[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷
H04N 7/18
H04N 5/00 H04N 5/76
H04N 5/91

[12] 发明专利申请公开说明书

[21] 申请号 01132807.X

[43] 公开日 2003年3月19日

[11] 公开号 CN 1404312A

[22] 申请日 2001.9.6 [21] 申请号 01132807.X

[71] 申请人 佳能株式会社
地址 日本东京
[72] 发明人 陈新武 纪 新

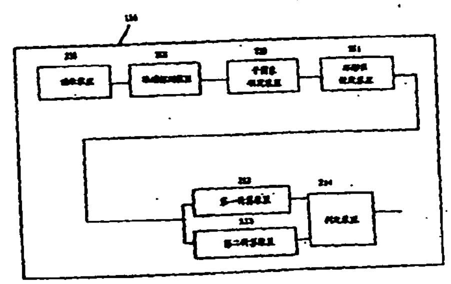
[74] 专利代理机构 中国国际贸易促进委员会专利商标事务所
代理人 李 强

权利要求书5页 说明书31页 附图28页

[54] 发明名称 图象处理方法和装置，图象处理系统以及存储介质

[57] 摘要

在此公布了用于探测图象中的人脸的方法、设备和系统。该方法包括：对于所述图象中的象素的一个子集，从所述图象的所述灰度分布导出一个第一变量；对于所述象素子集，从一个预设的基准分布导出一个第二变量，所述基准分布表征了所述对象；评价所述第一变量与所述第二变量在所述象素子集上的对应性；根据该评价步骤的结果判定所述图象是否包含所述对象。



ISSN1008-4274

知识产权出版社出版

Image Processing Methods and Apparatus for Identifying Human Eyes, Human Face, and Other Objects in an Image

Field of the Invention

The present invention relates to an image processing method, apparatus, and system for determining the human face in an image, and a storage medium.

Background of the Invention

Image processing method for detecting or extracting a feature region of a given image is very useful. For example, it can be used to determine the human face(s) in a given image. It is very useful to determine the human faces in an image, especially in an image with a complex background. Such a method can be used in many fields, such as telecommunication conferences, person-to-machine interface, security checking-up, monitor system for tracking human face, and image compression, etc.

It is easy for a human being (an adult or a baby) to identify human face in an image with a complex background. However, no efficient way has been found out to identify human face(s) in an image automatically and quickly.

Determining whether a region or a sub-image in an image contains a human face is an importing step in the human face recognition. At present, there are many ways for identifying human face. For example, a human face can be identified by making use of some salient features (such as two eyes, the mouth, the nose, etc.) and the inherent geometric positional relations among the salient features, or making use of the symmetric characters of

human face, complexion features of human face, template matching and neural network method, etc. For instance, a method is described in Haiyuan Wu, "Face Detection and Rotations Estimation using Color Information.", the 5th IEEE International Workshop on Robot and Human Communication, 1996, pp341-346, in which a method is given for utilizing human face features (two eyes and the mouth) and relations among the features to identify human face. In this method, the image region to be determined is first studied to find out whether the needed human face features can be extracted. If yes, then the matching degree of the extracted face human features to a known is human face model investigated, wherein the human face model describes the geometric relations among the human face features. If the matching degree is high, the image region is supposed to be an image of a human face. Otherwise, it is determined that the image region does not contains a human face. However, the method relies too much on the quality of the image to be investigated, and it is too much influenced by lighting conditions, the complexity of the image's background and the human race difference. Especially, it is very hard to determine human face exactly when the image quality is bad.

Summary of the Invention

Accordingly, it is an aim of the present invention to provide an improved image processing method, apparatus and storage medium for identifying objects in an image with a gray-level distribution.

It is a further aim of the present invention to provide an improved image processing method, apparatus, and storage medium. Said image processing method and apparatus can detect or identify an object in a given image quickly and effectively.

In the present invention, a method for determining an object in an image having a gray-level distribution is provided, which is characterized in that said method comprises the steps of:

selecting a subset of the pixels in said image;

for said subset of pixels, deriving a first variable from said gray-level distribution of said image,

for said subset of pixels, deriving a second variable from a preset reference gray-level distribution, said reference gray-level distribution being characteristic of said object;

evaluating the correspondence between said first variable and said second variable over the subset of pixels;

determining if said image contains said object based on the result of the evaluation step.

Further, in the present invention, a method for determining an object in an image having a gray-level distribution is provided, said method comprises the steps of:

a) determining a sub-image in said image;

b) selecting a subset of the pixels in said sub-image;

c) for said subset of pixels, deriving a first variable from said gray-level distribution of said image,

d) for said subset of pixels, deriving a second variable from a preset reference gray-level distribution, said reference gray-level distribution being characteristic of said object;

e) evaluating the correspondence between said first variable and said second variable over the subset of pixels;

f) determining if said image contains said object based on the result of the evaluation step.

In the present invention, object identification is carried out by evaluating two vector fields, so that adverse and uncertain effects such as non-uniform illumination can be eliminated, the requirement of the method of the present invention on the quality of the image is lowered, while the varieties of images that can be processed with the method of the present invention are increased. Moreover, gradient calculation is relatively simple, thus reducing the time needed for performing the identification.

As a specific embodiment of the invention, a sub-image, in which a target object (such as a human face) is to be identified, is determined by detecting one or more characteristic features (such as a pair of dark areas which are expected to be human eyes) in the image, providing an effective way for identifying a predetermined object(s) in the image.

As a specific embodiment, the method of the present invention limits the evaluation of correspondence between two vector fields in an area within the part of image to be detected, such as an annular area for human face detection, in which the correspondence between the two vector fields is more obvious, thus allowing the evaluation of the correspondence becoming more effective while reducing the time needed for carrying out the identification.

As a further embodiment, the method of the present invention performs a weighted statistical process during the evaluation, allowing the pixels with larger gradient (i.e. more obvious characteristic feature) have greater contribution to the outcome of the evaluation, thereby making the identification more effective.

As a further embodiment, the method of the present invention uses both weighted statistical process and non-weighted statistical process and determine if an object is included in the image to be detected on the basis of the results of both the processes, thus increasing the accuracy of the identification.

Also, the foregoing aim of the present invention is achieved by providing an image processing method for identifying an object in an image, comprising:

setting an annular region surrounding said rectangular region;

determining the gradient of gray level at each pixel in said annular region;

determining an ellipse for each pixel in said annular region, said ellipse passing through said pixel in said annular region, and determining the ellipse gradient of each pixel in the annular region; and

determining if said object is contained in said rectangle on the basis of the gradient of the gray level at each pixel and the ellipse gradient at each pixel in said annular region.

Further, the foregoing aim of the present invention is achieved by providing an image processing apparatus for determining a feature portion in an image, comprising:

means for determining a rectangular region to be identified in the image;

means for setting an annular region surround said rectangular region;

means for determining the gradient of gray level at each pixel in said annular region;

means for determining an ellipse for each pixel in said annular region, said ellipse passing through said pixel in said annular region, and for determining the ellipse gradient of each pixel in the annular region; and

means for determining if said object is contained in said rectangular region on the basis of the gradient of the gray level at each pixel and the ellipse gradient at each pixel in said annular region.

Moreover, the present invention provides a storage medium with a

program code for object-identification in an image with a gray-level distribution stored therein, characterized in that said program code comprises:

codes for determining a sub-image in said image;

codes for selecting a subset of the pixels in said sub-image;

codes for deriving, for said subset of pixels, a first variable from said gray-level distribution of said image,

codes for deriving, for said subset of pixels, a second variable from a preset reference gray-level distribution, said reference gray-level distribution being characteristic of said object;

codes for evaluating the correspondence between said first variable and said second variable over the subset of pixels; and

codes for determining if said image contains said object based on the result of the evaluation step."

Moreover, the present invention provides a human eye detection method for detecting human eyes in an image, comprising:

a read step of reading the gray level of each pixel in the each column in the image;

a segment step of segmenting each column into a plurality of intervals , and labeling each of the intervals as valley region , relay region or peak region;

a merging step of merging the valley region of the each column and the valley region of its adjacent column, and generating an eye candidate region; and

a determination step of determining the human eye from the eye candidate regions.

The other objects and features of the present invention will become apparent from the following embodiments and drawings. The same reference

numeral in the drawings indicates the same or the like component.

Brief Description on the Drawings

The accompanying drawings, which are incorporated herein and constitute a part of the specification, illustrate embodiments of the invention and, together with the description, serve to explain the invention.

Fig. 1 is a block diagram showing an image processing system of the present invention.

Fig. 2 is a block diagram showing an arrangement of human face identification apparatus according to the present invention.

Fig. 3 schematically shows an original image to be detected.

Fig. 4 is a flow chart showing a human face identifying process according to the first embodiment of the present invention.

Fig. 5 is a diagram showing a rectangular region of Fig. 3 and the determined annular region around it.

Fig. 6 is a diagram showing several pixels in an image.

Fig. 7 is a diagram showing how to determine the ellipse gradient for a pixel in the annular region.

Fig. 8A is a diagram showing another example of an original image to be detected.

Fig. 8B is a diagram showing a rectangular region in the original image of Fig. 8A.

Fig. 8C is a diagram showing a annular region determined for the rectangular region in Fig. 8B.

Fig. 8D is a diagram showing another rectangular region in the original image in Fig. 8A.

Fig. 8E is a diagram showing a annular region determined for the rectangular region in Fig. 8D.

Fig. 9 is a flow chart showing the human face identification process of another embodiment according to the present invention.

Fig. 10 shows a reference gray-level distribution for use in identifying human face in an image;

Fig. 11 is for showing a way for generating a sub-image for human face detection from a pair of dark (eye) areas detected, which are expected to correspond a pair of human eyes.

Fig. 12 is a block diagram showing the arrangement of an eye detection device according to an embodiment of the present invention;

Fig. 13A is a flow chart showing the procedure of searching human eye areas.

Fig. 13B is an example of an original image to be detected.

Fig. 14A is a flow chart for segmenting every each column in an image.

Fig. 14B is an example for showing a column of pixels in an image.

Fig. 14C is an example for showing the gray level distribution of a column.

Fig. 14D is a diagram showing the gray level of a column segmented into intervals.

Fig. 14E is an example for showing a segmented column in an image.

Fig. 14F is a diagram showing the determination of a segment point in a column.

Fig. 15A is a flow chart showing the process for merging valley regions in the columns.

Fig. 15B is a diagram showing the columns in an image and the valley regions and the seed regions in each column.

Fig. 15C is an image showing the detected candidate eye areas.

Fig. 16A is the flow chart showing the process for determining eye areas in accordance with the present invention.

Fig. 16B is a diagram showing a candidate eye area and its circum-rectangle.

Fig. 16C is an image showing the detected eye areas.

Fig. 17A is a flow chart showing the process for adjusting segment border.

Fig. 17B is a diagram showing the merger of a segment point to its adjacent intervals.

Fig. 17C is a diagram showing the merger of an intermediate region to its adjacent valley region.

Fig. 18A is a flow chart showing a process for judging whether a valley region can be merged into a seed region.

Fig. 18B is a diagram showing a seed region's predicted valley region.

Fig. 18C is a diagram showing an overlap between two valley regions.

Description of the Preferred Embodiments

Preferred embodiments of the present invention will now be described in detail with reference to the accompanying drawings.

Fig. 1 is a block diagram showing an image processing system utilizing an image processing apparatus of the present invention. In the system, a printer 105, such as an ink-jet printer or the like, and a monitor 106 are connected to a host computer 100.

Operating on the host computer 100 are an application software program 101 such as a word-processor, spreadsheet, Internet browser, and the like, an OS (Operating System) 102, a printer driver 103 for processing various drawing commands (image drawing command, text drawing command, graphics drawing command) for instructing the output of images, which are issued by the application software program 101 to the OS 102 for generating print data, and a monitor driver 104 for processing various drawing commands issued by the application software program 101 and displaying data in the monitor 106.

Reference numeral 112 denotes an instruction input device; and 113, its device driver. For example, a mouse that allows a user to point to and click on various kinds of information displayed on the monitor 106 to issue various instructions to the OS 102 is connected. Note that other pointing devices, such as trackball, pen, touch panel, and the like, or a keyboard may be connected in place of the mouse.

The host computer 100 comprises, as various kinds of hardware that can ran these software programs, a central processing unit (CPU) 108, a hard disk

(HD) 107, a random-access memory (RAM) 109, a read-only memory (ROM) 110, and the like.

An example of the face identification system shown in Fig. 1 may comprises Windows 98 available from Microsoft Corp. installed as an OS in a PC-AT compatible personal computer available from IBM Corp., desired application program(s) installed that can implement printing and a monitor, and a printer connected to the personal computer.

In the host computer 100, each application software program 101 generates output image data using text data (such as characters or the like), graphics data, image data, and so forth. Upon printing out the output image data, the application software program 101 sends a print-out request to the OS 102. At this time, the application software program 101 issues a drawing command group that includes a graphics drawing command corresponding to graphics data and an image drawing command corresponding to image data to the OS 102.

Upon receiving the output request from the application software program 101, the OS 102 issues a drawing command group to the printer driver 103 corresponding to an output printer. The printer driver 103 processes the print request and drawing commands inputted from the OS 102, generates print data for the printer 105, and transfers the print data to the printer 105. The printer driver 103 performs an image correction process for the drawing commands from OS 102, and then rasterizes the commands sequentially on a memory, such as a RGB 24-bit page memory. Upon completion of rasterization of all the drawing command, the printer driver 103 converts the contents of the memory into a data format with which the printer 105 can perform printing, e. g., CMYK data, and transfers the converted data

to the printer 105.

Note that the host computer 100 can connect a digital camera 111, that senses an object image and generates RGB image data, and can load and store the sensed image data in the HD 107. Note that the image data sensed by the digital camera 111 is encoded, for example by JPEG. The sensed image data can be transferred as image data to the printer 105 after it is decoded by the printer driver 103.

The host computer 100 further comprises a face identification apparatus 114 for determining the human face in an image. The image data stored in HD 107 are read and processed by the face identification apparatus 114. First, the possible positions of the human face region are determined and read, and whether the region contains one or more human faces is determined. Then, the portion(s) of the image that contains the determined human face(s) in the image can be sent to the printer 105 or monitor 106 under the control of OS 102.

Face identification Apparatus

Fig. 2 is a block diagram showing the arrangement of the face identification apparatus according to the present invention.

The face identification apparatus 114 of the present embodiment comprises a reading means 210, an eye area detecting device 218, a sub-image determining device 219, an annular region setting means 211, a first determination means 212, a second determination means 213, and a third determination means 214. In the face identification apparatus 114, a reading means 210 executes an image reading process. The gray level of each pixel of an image stored in the HD 107 or the RAM 109 or the like is read by the

reading means 210.

Human Face Identification Process

Fig. 3 schematically shows an example of an original image to be detected. The original image contains a human face. The original image 300 can be input to the human face identification system by a digital device 111 such as a digital camera, a scanner or the like, and the original image is stored in the HD 107 or the RAM 109, or the like.

Referring to Fig. 3, it can be seen that the shape of the human face contour in the original image 300 is generally closed to an ellipse, which has nothing to do with human race, complexion, age and gender. Along the human face contour, a circumscribed rectangular region 300A can be drawn.

Fig. 4 is a flow chart for showing a human face identification process according to an embodiment of the present invention.

Referring to the flow charts in Fig. 4 and Fig. 3, an explanation to human face identification process for the original image will be given.

Reading Original Image and Determining Sub-Image (Rectangular region) to be Identified

Referring to Fig. 4, the human face identification process starts in step S40. In step S41, reading means 210 reads an original image 300 to be detected, and acquires the gray level of each pixel of the original image 300. If the original image 300 is encoded by, e.g., JPEG, the reading means 210 must first decode it before reads its image data. Of course, all the gray levels for every columns of the original image 300 can be read simultaneously, or they can be read sequentially or respectively. In step S41, the sub-image

determining device 218 determines one or more sub-images (or regions) 300A in the original image 300 for human face identification, and it also determines the location of the sub-image 300A in the original image 300. The sub-image 300A can be substantially rectangular and is the candidate region of a human face image portion. However, the shape of the sub-image is not limited to rectangular but can be any other suitable shape.

A method and device, in accordance with the present invention, for determining sub-image(s) 300A in an image for identification are described below. It is to be noted, however, that the manner of eye area detection for determining the sub-image is not limited to the method of the present invention as described below. Instead, other methods and/or processes, as those known in the art, can be utilized for determining the sub-image.

Determination of Sub-image 300A

Eye Detection Device

Fig. 12 is a block diagram showing the arrangement of the eye detection device according to the an embodiment of the present invention;

The eye detection device 1114 of the present embodiment comprises a reading means 1200, a segment means 1201, a merger means 1202 and a determination means 1203. In the eye detection device 1114, a the reading means 1200 performs an image-reading process , in which the gray level of each pixel in the columns of an image stored in, for example, HD 1107 is read by the reading means 1200. Referring to Figs. 14D and 14E, on the basis of the gray level of each pixel in a column C41 of an image, the column C41 of an image is segmented into a plurality of intervals I1-1, I1-2, ... I1-9, I1-10 by segment means 1201. These intervals I1-1, I1-2, ... I1-9, I1-10 can be marked as classified into three types: peak regions, valley regions and

intermediate regions, according to their average gray level of the image data. The terms "peak region, valley region and intermediate region" will be defined in detail later. Then, the valley regions of column C41 can be obtained. In the same way, the segment means 1201 also divides other columns of the image into the three types and obtains their valley regions respectively. After all the columns of an image have been marked as the three types and their valley regions have been obtained, the merger means 1202 executes the merging process and merges the valley regions in the adjacent columns. The merged valley regions are set as the human eye candidates. Then the human eye can be determined by determination means 1203.

Detecting the eye areas

A human eye detection process for an original image will be explained below with reference to the flow chart in Fig. 13A. Fig. 13B is an example of an original image to be detected. Assume that the original image is stored in a predetermined area in the HD 1107 or the RAM 1109, or the like.

Referring to Fig. 13A, in step S131, the reading means 1200 reads the gray level of each pixel in the columns of the original image to be detected. If the original image is encoded by, e.g., JPEG, the reading means 1200 must first decode it before reads its image data. Of course, all the gray levels for every columns of the original image can be read at one time, or they can be read sequentially or respectively. In the preferred embodiment, each column of the image is read, . However, each row of the image can be read, as in case the given image has been related to 90°. In step S132, every columns of the original image are is segmented into many intervals by the segment means 1201. With reference to Fig. 14E, the lengths of each of the intervals I1-1, I1-2, ... I1-9 and I1-10 is variable. For example, the length of interval I1-1 is not equal to the length of interval I1-2. Some of the segmented intervals are marked as the valley regions on the basis of their average gray levels of pixels.

In step S133, the valley regions in the adjacent columns are merged by the merger means 1202 to generate the eye candidate regions. Since the valley regions in each column have different lengths, the sizes of the eye candidate regions are also different from each other. In step S134, the human eye areas in the eye candidate regions are determined by the determination means 1203. Thus, areas corresponding to human eyes in the image can be detected.

Segmenting Each Column of an Image

Fig. 14A is the flow chart showing the process for segmenting each column in an image in step S132.

The terms " valley region", "peak region" and "intermediate region" are defined as below.

Fig. 14B is an example for showing a column in the image. Referring to Fig. 14B, a column C41 of the original image is read by the reading means 1200. Fig. 14C shows a gray level distribution of the column C41. Fig. 14D is a gray level distribution of the column segmented into intervals. In Fig. 14D, the reference numerals I1-5, I1-6, I1-9 denote the segmented intervals, respectively, and the gray level of each of the segments or intervals is the average of the gray levels of pixels in the same segment or interval in Fig. 14C.

Fig. 14E is the segmented column of the image in Fig. 14B. Referring to Fig. 14E, the image data of a column C41 in the image is read by reading means 1200. For the image of Fig. 14B, the column C41 is segmented into 10 intervals I1-1, I1-2, ... I1-9 and I1-10. An interval's size is the number of the pixels in the interval. For example, if the interval I1-2 comprises 12 pixels, the interval I1-2's size is 12.

With reference to Fig. 14D and 14E, if an interval's gray level is less than both of its adjacent intervals' gray level, then the interval is called a valley region. If an interval's (average) gray level is bigger than both of its adjacent

intervals' gray level, the interval is called a peak region. On the other hand, if an interval's gray level is between its adjacent interval's gray levels, such an interval is called an intermediate region. As to column C41 of the embodiment, the gray levels of intervals from I1-1 to I1-10 are 196, 189, 190, 185, 201, 194, 213, 178, 188, and 231 respectively. As to interval I1-6, its gray level is 194, and the gray levels of its adjacent intervals I1-5 and I1-7 are 201 and 213 respectively. Since the gray level of interval I1-6 is less than that of its adjacent intervals I1-5 and I1-7, the interval I1-6 is determined as a valley region. In the same way, intervals I1-2, I1-4 and I1-8 are also determined as valley regions. As to interval I1-5, its gray level is 201, and the gray levels of its adjacent intervals are 185 and 194 respectively. Since the gray level of interval I1-5 is bigger than that of its adjacent intervals I1-6 and I1-7, the interval I1-5 is determined as a peak region. In the same way, intervals I1-1,I1-3,I1-7 and I1-10 are also determined as peak regions. Further, as to interval I1-9, its gray level is 188, the gray levels of its adjacent I1-8 and I1-10 are 178 and 231. Since the gray level of interval I1-9 is between the gray levels of its adjacent intervals I1-8 and I1-10, interval I1-9 is determined as an intermediate region.

As a valley region is also an interval, the ways for computing the valley region's gray level and size are the same as those for computing the interval's gray level and size. It is also applied to the computation of the gray level and size of a peak region or an intermediate region.

The process for segmenting every column in an image in step S132 will be explained below with reference to Fig. 14A.

Referring to Fig. 14A, the gray level of each pixel in the first column from the left of the detected image are read out in step S141. In order to segment the column into intervals of the three types, i.e., valley regions, peak regions and intermediate regions, the segmenting points have to be determined.

In step S142, whether a pixel in the column is a segment point can be determined according to the values of first and second-order derivatives of the gray-level distribution at the pixel. Fig. 14F is a diagram for showing the procedure to determine whether a pixel is a segmenting point in a column. With reference to Fig. 14F, two adjacent pixels Pil and Pi2 are given in a column. Then the values of the first and the second-order derivatives at these two pixels Pi1, Pi2 can be easily obtained by the following formula:

$$F(x) = f(x) \times g(x) = \int_{-\infty}^{+\infty} f(t) \cdot g(x-t)dt.$$

wherein, x is the position of a pixel, f(x) is the gray level of the pixel at x, g(x) is Gaussian function, i.e. $g(x) = \exp(-x^2/2)\sqrt{2}$. F(x) is the convolution of f(x) and g(x). Thus, the values of the first-order derivative and second order derivative can be determined as the values of the first-order and second-order derivative of the gray level distribution at the pixel.

As an example, the first-order derivative values of the pixels Pil and Pi2 are assumed as D1f and D2f, the second-order derivative values of the pixels Pil and Pi2 are assumed as D1s and D2s, where D1f is the first-order derivative value of pixel Pi1, D2f is the first-order derivative value of pixel Pi2, D1s is the second-order derivative value of pixel Pi1, and D2s is the second-order derivative value of pixel Pi2. If the product of D1s and D2s is less than zero, or D2s is equal to zero, and the absolute value of D2f is bigger than a predetermined value, then the pixel Pi2 is determined as a segmenting point. Otherwise, the pixel Pi2 is determined as a non-segmenting point.

Thus, the segmenting points s11, s12, ... s19 can be obtained in step S142.

After the segmenting points in a column have been determined, the column can be segmented into a plurality of intervals in step S143. Then, in

step S144, the intervals are divided into valley region, peak region and intermediate region in accordance with the gray levels of the intervals. The border of intervals is adjusted in step S145. The detail of step S145 will be described using detailed flow chart. In step 146, it is checked if all columns in the detected image have been segmented. If the column being segmented is not the last column of the detected image, the flow goes to step S147. In step S147, the gray levels of pixels in the next column are read out. Then the flow returns to step S142 to repeat the process in step S142 and the subsequent steps. However, if the column being segmented is the last column of the detected image in step 146, i.e. all columns have been segmented, the flow ends in step S148.

Alternatively, the above-mentioned segmenting process may start from the first column from the right of the detected image.

Merging Valley Regions to Generate Eye Candidate Regions

Fig. 15A is the flow chart for showing the process of merging valley region in the columns in step S133 in Fig. 13A. Fig. 15B is a diagram showing the columns of an image and the valley regions and seed regions in each column of the image. In Fig. 15B, an image has n columns Col1, Col2, ... Coln.

With reference to Fig. 15A and 15B, all the valley regions S1, S2, S3 and S4 in the first column Col1 (most left) of the detected image are set as seed regions in step S151. A seed region is an aggregation of one or more valley regions. Since the gray level of a valley region is less than that of a peak region or a intermediate region, a seed region is usually a dark area in a column.

In step S152 of Fig. 15A, the first valley region V2-1 in the next column Col2 is read out. Then the flow advances to step S153. In step S153, the first seed region S1 is read out. In step S154, it is checked if the valley region V2-

1 of column Col2 can be merged into the seed region S1 on the basis of the valley region V2-1 and the seed region S1. If the valley region V2-1 of the valley region V2-1 can be merged into the seed region S1, then the flow goes to step S156 and the valley region V2-1is merged into the seed region, then the valley region becomes a part of the seed region. However, if it is checked in step S154 that the valley region V2-1 can not be merged into the seed region S1, the flow goes to step S155. In the present example, valley region V2-1 of column Col2 can not be merged into seed region S1. The flow advances to step S155. In step S155, it is checked whether the seed region is the last seed region. If the seed region is not the last seed region, then next seed region is read out in step S157 and the flow returns to step S154 to repeat the processes in step S154 and the subsequent steps. In the present example, seed region S1 is not the last seed region, so in step S157 the next seed region S2 is read out, and the steps of S154 and S155 are repeated. If it is checked in step S155 that the seed region is the last seed region (for example, the seed region S4 as shown in Fig 5B), then the flow advances to step S158 and set the valley region that can not be merged into a seed region as a new seed region. Referring to Fig. 15B, since valley region V2-1 of column Col2 can not be merged into seed regions S1, S2, S3 or S4, that is, it is a valley region that can not be merged into any existing seed region, then valley region V2-1 of column Col2 is set as a new seed region in step S158.

In step S159, it is checked if all the valley regions in the column Col2 have been processed. If all the valley regions in the column Col2 have been processed, the flow goes to step S1511. In step S1511, it is checked if all the columns have been processed. If the column is not the last column of the detected image, then the flow returns to step S152 to repeat the processes in step S154 and the subsequent steps. As column Col2 is not the last column of the detected image, the flow returns to step S152. If all the columns have been processed, i.e., if the column is the last column Coln, the flow advances to

step S1520. In step S1520, all the seed regions are set as eye candidate regions. Then the flow ends in step S1521. Fig. 15C is an example showing the result for merging valley regions to generate eye candidate regions in columns in a detected image in step S133.

Determining Eye Areas

Fig. 16A is the flow chart showing the process for determining eye areas in step S134.

With reference to Fig. 16A, the first eye candidate region is read out in step S161. Then, the flow advances to step S162. In step S162, the gray level of an eye candidate region is calculated. As described above, an eye candidate region comprises one or more valley regions. If an eye candidate region is comprised of n valley regions, i.e. valley region 1, valley region 2, ... valley region n, then the eye candidate region's gray level calculated in step S162 is given by :

$$EyeGrayl = (Valley1Grayl + Valley2Grayl ... + ValleynGrayl)/n \quad (1)$$

where

EyeGrayl is an eye candidate region's gray level;

Valley1Grayl is the gray level of valley region 1;

Valley2Grayl is the gray level of valley region 2...

ValleynGrayl is the gray level of valley region n; and

n is the number of valley regions included in an eye candidate region.

Therefore, if an eye candidate region comprises 3 valley regions with gray levels of 10, 20 and 30, respectively, then the gray level of the eye candidate region will be (10+20+30)/3=20.

Referring to step S162 of Fig. 16A, the gray level of an eye candidate region is calculated. If the eye candidate region's gray level is not less than a

first threshold (for example, 160), the flow goes to step S1610. In the present embodiment, the first threshold is within the range of 100 to 200. In step S1610, the eye candidate region is determined as a false eye area and is rejected. Then the flow goes to step S168. In step S168, it is checked if all the eye candidate regions in the detected image have been processed. If it is not the last eye candidate region, then the next eye candidate region will be read in step S169, then the flow returns to step S162 to repeat the processes in step S162 and the subsequent steps. However, if it is checked in step S168 that the eye candidate region is the last eye candidate region, then all eye candidate regions in the detected image have been determined, and the flow ends in step S1611.

At step S162, if the gray level of the eye candidate region is less than the first threshold, the flow advances to step S163.

The background gray level of the eye candidate region is calculated in step S163. The background gray levels of valley regions included in the eye candidate region determine the background gray levels of an eye candidate region. A valley region's background gray level is the average gray level of its adjacent intervals' gray levels. The eye candidate region's background gray level calculated in step S163 is given by :

EyeBGrayl = (Valley1BGrayl +Valley2BGrayl...+ValleynBGray l)/n;
(2)

where
EyeBGaryl is an eye candidate region's background gray level;
Valley1BGrayl is the background gray level of valley region 1;
Valley2BGray1 is the background gray level of valley region 2...
ValleynBGrayl is the background gray level of valley region n; and
n is the number of valley regions included in an eye candidate region.

At step S163, the background gray level of an eye candidate region is calculated. If the eye candidate region's background gray level is not bigger than a second threshold(for example, 30) in step S163, the flow goes to step S1610. In the present embodiment, the second threshold is within the range of 20 to 80. For the present example, in step S1610, the eye candidate region is determined as a false eye area and is rejected. Then the flow goes to step S168.

At step S163, if the background gray level of the eye candidate region is bigger than the second threshold, the flow advances to step S164.

The difference of background gray level of the eye candidate region and the gray level of the eye candidate region is calculated in step S164. If the difference is not bigger than a third threshold(for example, 20), the flow goes to step S1610. In the present embodiment, the third threshold is within the range of 5 to 120. In step S1610, the eye candidate region is determined as a false eye area and is rejected. Then the flow goes to step S168.

At step S163, if the difference of background gray level of the eye candidate region and gray level of the eye candidate region is bigger than the third threshold, the flow advances to step S165.

The ratio of the width to the height of an eye candidate region is calculated in step S165.

As to the height and the width of an eye candidate region, we have the following definitions. Valley region's size is the number of pixels included in a valley region. For example, if a valley region comprises 5 pixels, then the valley region's size is 5. The size of an eye candidate region is the sum of the sizes of the valley regions included in the eye candidate region. The width of an eye candidate region is the number of valley regions included in the eye candidate region. The height Hd of an eye candidate region is given by:

$$Hd = Sd/Wd \qquad (3)$$

Where,

Hd is the height of an eye candidate region;

Sd is the size of an eye candidate region;

Wd is the width of an eye candidate region.

With reference to step S165 in Fig. 16A, the ratio of the width to the height of an eye candidate region is calculated. If in step S165 the ratio of the width to the height of an eye candidate region is not bigger than a fourth threshold(for example, 3.33), the flow goes to step S1610. In the present embodiment, the fourth threshold is within the range of 1 to 5. In step S1610, the eye candidate region is determined as a false eye area and is rejected. Then the flow goes to step S168.

At step S165, if the ratio of the width to the height of an eye candidate region is bigger than the fourth threshold, the flow advances to step S166.

The ratio of the size of an eye candidate region to that of its circum-rectangle is calculated in step S166. Fig. 16B is a diagram showing an eye candidate region and its circum-rectangle. With reference to Fig. 16B, an eye candidate region D1 and its circum-rectangle DC1 are given. As seen from Fig. 16B, the eye candidate region's circum-rectangle DC1 is the smallest rectangle that encircles the eye candidate region D1. The size of an eye candidate region's circum-rectangle is the number of pixels included in the circum-rectangle. The size of an eye candidate region is the number of pixels included in the eye candidate region.

At step S166, the ratio of the size of an eye candidate region to that of its circum-rectangle is calculated. If the ratio is not bigger than a fifth threshold (for example 0.4) in step S166, the flow goes to step S1610. In the present embodiment, the fifth threshold is within the range of 0.2 to 1. In step S1610, the eye candidate region is determined as a false eye area and is rejected. Then the flow goes to step S168.

At step S166, if the ratio of the size of an eye candidate region to that of its circum-rectangle is bigger than the fifth threshold, the flow advances to step S167, where the eye candidate region is determined as a true eye area.

After step S167, the flow advances to step S168 and determines whether the eye candidate region is the last eye candidate region. If NO, then the next eye candidate region is read in step S169 and the flow returns to step S162. If YES in step S168, then all eye areas are determined. Fig. 16C is an example showing the result for detecting eye areas in an image in step S133.

Adjusting Segment's Border

Fig. 17A is a flow chart showing the process of adjusting segment border in step S145 in Fig. 14A.

With reference to Fig. 17A, the gray level of a segment point is compared with the gray levels of its two adjacent intervals, then the point is merged into the interval whose gray level is closer to the point's gray level in step S171. For example, referring to Fig. 17B, the gray level of segment point S is 80, its adjacent intervals are intervals In1 and In2. The gray levels of intervals In1 and In2 are 70 and 100 respectively. Since the gray level of interval In1 is closer to that of point S, then S is merged into interval In1.

Further, the flow advances to step S172. In step S172, the first intermediate region is read out. Then the gray levels of the intermediate region and its adjacent valley region and peak region are calculated in step S173. After the gray levels of them have been calculated, the flow advances to step S174. In step S174, a comparison is made to decide whether GR is less than $GP \times Th6 + GV \times (1 - Th6)$, wherein, GR denotes the gray level of the intermediate region, GV denotes the gray level of the intermediate region's adjacent valley region, GP denotes the gray level of the intermediate region's adjacent peak region. Th6 is the sixth threshold(for example, 0.2). In the present embodiment, the sixth threshold is within the range of 0 to 0.5. If the

decision of step S174 is NO, the flow goes to step S176. If the decision of step S174 is YES, then the intermediate region is merged into the valley region in step S175.

Fig. 17C is a diagram showing an example for merging the intermediate region to the adjacent valley region. X axis shown in Fig. 17C represents the position of each column, Y axis shown in Fig. 17C represents the gray level of each region.

Referring to Fig. 17C, intermediate region Re1's gray level is 25, valley region Va1's gray level is 20, and peak region Pe1's gray level is 70. When the sixth threshold is set as 0.2, then

$$GP \times Th6 + GV \times (1 - Th6) = 70 \times 0.2 + 20 \times 0.8 = 30 > GR = 25$$

Therefore, the decision in step S174 is YES, so the intermediate region Re1 will be merged into the valley region Va1. Further, intermediate region Re2's gray level is 40,peak region Pe2's gray level is 60, then

$$GP \times Th6 + GV \times (1 - Th6) = 60 \times 0.2 + 20 \times 0.8 = 28 < GR = 40;$$

Therefore, the decision in step S174 is NO, so intermediate region Re2 will not be merged into the valley region Va1.

Referring to step S176 in Fig. 17A, it is checked if all the intermediate regions in the detected image have been processed. If the intermediate region is not the last intermediate region, the next intermediate region will be read in step S177, then the flow returns to step S173 to repeat the processes in step S173 and the subsequent steps. However, if it is checked in step S176 that the intermediate region is the last intermediate region, i.e. the processing for all intermediate regions is complete, the flow ends in step S178. Thus, all segment borders in the detected image have been adjusted.

determining Whether a Valley Region Can be Merged Into a Seed Region

Fig. 18A is the flow chart for showing the process for determining whether a valley region can be merged into a seed region in step S154 in Fig.

15A.

Fig. 18B is a diagram showing a seed region's predicted valley region. A seed region's predicted valley region isn't a real existing valley region in any columns of the detected image. It is a valley region that is assumed in the next column of the column including the most adjacent valley region at the right of the seed region, and its position is the same as that of the most adjacent valley region at the right of the seed region. With reference to Fig. 18B, valley region Va3 is the most adjacent valley region at the right of seed region Se1. Valley region Va3 is in the column Col1, and column Col2 is the next column of column Col1. Then valley region Va1 is the predicted valley region of seed region Se1. This predicted valley region is in column Col2, and its position is the same as that of valley region Va3 but in a different column.

Fig. 18C is a diagram showing an overlap region of two valley regions. The overlap region of two valley regions is an area in which the pixels belong to the two valley regions.

Referring to Fig. 18C, the interval from point B to point D is a valley region Va1, the interval from point A to point C is a valley region Va2, the valley region Va1 is the predicted valley region of the seed region Se1, the valley region Va2 is a real valley region in column Col2. Then, the interval from point B to point C is the overlap region of the valley region Va1 and the valley region Va2.

The procedure for judging whether a valley region can be merged into a seed region will be explained below with reference to the flow chart in Fig. 18A. Referring to Fig. 18A, the overlap region of a valley region and a seed region's predicted valley region is calculated in step S181.

After the overlap region has been calculated, the flow advances to step S182. In step S182, a comparison is made to decide whether Osize/Max(Vsize, SVsize) is bigger than Th7, wherein, Osize is the size of overlap of the valley region and the seed region's predicted valley region, Max (Vsize, SVsize) is

the maximum of the size of the valley region and that of the seed region's predicted valley region, and Th7 is the seventh threshold (for example, 0.37). The seventh threshold is within the range of 0.2 to 0.75.

If the decision of step S182 is NO, the flow goes to step S188; then the valley region can not be merged into the seed region, and the flow ends in step S189. Otherwise, if the decision of step S182 is YES, then the flow advances to step S183.

In step S183, the gray levels of the valley region and the seed region are calculated. Then the flow advances to step S184. In step S184, it is determined whether |GValley-GSeed| is less than Th8, wherein GValley is the gray level of the valley region, GSeed is the gray level of the seed region, and Th8 is an eighth threshold(for example, 40). The eighth threshold is within the range of 0 to 60. If the decision of step S184 is NO, the flow goes to step S188; then the valley region can not be merged into the seed region, and the flow ends in step S189. Otherwise, if the decision of step S184 is YES, then the flow advances to step S185.

In step S185, the luminance values of the valley region's background, the seed region's background, the valley region and the seed region are calculated.

As to the luminance value of a pixel in an image, it can be calculated by:

$$G=0.12219\times L-0.0009063\times L^2+0.000036833526\times L^3-0.0000001267023\times L^4+0.0000000001987583\times L^5; \qquad (4)$$

where G is gray level of a pixel ranged from 0 to 255; and L is luminance level of a pixel and is also ranged from 0 to 255.

Therefore, the luminance value of a pixel can be obtained from its gray level in an image. On the other hand, the gray level of a pixel can be obtained from its luminance value.

For the present example, the pixels Pi1 and Pi2 in Fig. 14 F have the gray levels of 50 and 150 respectively. With formula (4), it can be determined that the luminance valves of pixels Pi1 and Pi2 are 128 and 206 respectively.

Referring to Fig. 18A, after step S185, the flow advances to step S186. In step S186, it is determined whether Min((Lvb-Lv), (Lsb-Ls))/Max((Lvb-Lv), (Lsb-Ls)) is bigger than Th9, wherein Lv is the luminance value of the valley region, Ls is the luminance value of the seed region, Lvb is the luminance value of the valley region's background, Lsb is the luminance value of the seed region's background. Min ((Lvb-Lv), (Lsb-Ls)) is the minimum of (Lvb-Lv) and (Lsb-Ls), Max ((Lvb-Lv), (Lsb-Ls)) is the maximum of (Lvb-Lv) and (Lsb-Ls), and Th9 is a ninth threshold (for example, 0.58). The ninth threshold is within the range of 0.3 to 1. If the decision of step S186 is NO, the flow goes to step S188; then the valley region can not be merged into the seed region, and the flow ends in step S189. Otherwise, if the decision of step S186 is YES, the flow advances to step S187.

In step S187, the valley region is merged into the seed region, then the flow ends in step S189.

As can be seen from the above, the present invention provides a method and a device for quickly detecting areas, each of which corresponds to a human eye, in an image with a complex background, without requiring that detected image has a very high quality, thus greatly reducing the possibility that human eyes in the image fails to be detected. The method of the invention allows for the precision detection of human eyes under different scales, orientation and lighting condition. Therefore, with the method and device of the present invention, the human eyes in an image can be quickly and effectively detected.

The method and device of the present invention explained above are described with reference to detection human eyes in an image, however, the method and device of the present invention are not limited to detection human

eyes, it is applicable to other detection method, for example, method to detect flaw portion on a circuit board.

Moreover, the shape of the sub-image 300A for identification is not limited to rectangular, rather, it can be any suitable shape, e.g. ellipse etc.

In addition, the above method and device for detecting areas corresponding to human eyes in an image are used for identification of human face contained in an image, but the method and apparatus for the identification of human face in an image may also use other methods and devices for detection of human eyes in the image on which identification is to be performed, for example, the method disclosed in Kin-Man Lam, "A Fast Approach for Detecting Human faces in a Complex Background ", Proceedings of the 1998 IEEE International Symposium on Circuits and System, 1998, ISCAS'98 Vol. 4, pp85-88 can be used in the method and apparatus of the present invention for human face identification.

Upon detecting at least two eye (or dark) areas, an arbitrary pair of the eye areas is selected as a pair of candidate eyes. For each pair of eye areas, the distance L between the centers of them is determined. Then, a rectangular sub-image 300A is determined as show in Fig. 11. For human face detection, in addition to the rectangular sub-image as shown in Figure 11, another rectangular sub-image is also determined for detection of human face, which is symmetrical to the rectangular sub-image of Figure 11 with respect to the lines passing the centers of the pair of eye areas.

It is to be noted that the values/ratios shown in Fig. 11 are not necessarily strict, rather, values/ratios within certain range (such as ±20 or so) of those as shown in Fig. 11 are acceptable for performing the method for human face identification of the present invention.

In the case that more than two eye areas are detected, each possible pair of eye areas are chosen for determination of respective sub-images. And for each pair of eye areas in the image, two rectangular sub-images may be determined in the manner as described above.

Then, for each sub-image thus determined, identification is performed as to whether the sub-image corresponds to a picture of human face, as described below.

Determination of Annular region

Then, the flow goes to step S42. In step S42, a annular region 300R (Fig. 5) surrounding the sub-image (rectangular region) 300A is determined by a annular region determination means 211. The generation of the annular region 300R will be described in detail with reference to Fig. 5.

Fig. 5 is a diagram showing the rectangular region 300A of Fig. 3 and the determined annular region 300R around it. In the Cartesian coordinate system shown in Fig. 5, the upper left corner of the original image 300 is taken as the coordinate origin, wherein X and Y axes extend in horizontal or vertical directions respectively.

Referring to Fig. 5, by making use of reading means 210 as shown in Fig. 2, the pre-stored original image 300, as shown in Fig. 3, is read from the HD 107 or the RAM 109, and the location or rectangular region 300A relative to the original image 300 is also acquired. In the present embodiment, the coordinates of each of the four corners of rectangular region 300 are (230, 212), (370, 212), (230, 387) and (370, 387) respectively in the given Cartesian coordinate system, thus its width and length is 140 and 175 respectively.

From rectangle 300A, two rectangles 300A1 and 300A2 are derived,

Rectangle 300A1 is linearly enlarged with respect to rectangle 300A with a factor of 9/7, and rectangle 300A2 is reduced with respect to rectangle 300A with a factor of 5/7. Two rectangular regions 300A1 and 300A2 are produced. Thus the coordinate of each of the four corners of the first rectangular region 300A1 in this coordinate system is (210, 187), (390, 187), (210, 412) and (390, 412)respectively; while the coordinate of each of the four corners of the second rectangular region 300A2 in this coordinate system is (250, 237), (370, 237), (250, 362) and (370, 362) respectively. The region between the first rectangular region 300A1 and the second 300A2 forms the annular region 300R for the above-mentioned rectangular region 300A.

The annular region 300R shown in Fig. 5 is produced by first enlarging the width and length of the rectangular region 300A in a factor of 9/7 and then reducing that in a factor of 5/7. But the present invention is not limited to it, as the width and length of the larger rectangular region can be at a factor of n and those of the smaller rectangle at a factor of m, thus the larger rectangular region and the smaller 300A2 rectangular region are respectively produced, where m within the range of 0 to 1 and n within the range of 1 to 2.

Determining the Gradient of the Gray Level at Each Pixel in the Annular region and its Weight

Returning to Fig. 4, the flow goes to step S43 after step S42. In step S43, the gradient of the gray level at each pixel in the annular region 300R is determined by a first determination means 212, while the weight of the gradient of gray level at each pixel in the annular region 300R is also determined.

With reference to Fig. 5 and Fig. 6, how the human face identification

apparatus according to the present invention determines the gradient and its corresponding weight of the gray level at each pixel in the annular region 300R is explained.

In general, each pixel in an image is surrounded by many other pixels. Hence, the gradient of the gray level of a pixel can be determined by the gray levels of k^2 pixels around the certain pixel. Where, k is an integer ranged from 2 to 15. In the present embodiment, k is given to be 3. In the present invention, the Sobel operator is used to calculate the gradient of the gray level distribution at each pixel in the annular region 300R in the original image.

Referring to Fig. 6, how to determine the gradient of the gray level of each pixel in the annular region 300R is explained, wherein the coordinates of pixel P in the coordinate system is (380, 250).

Referring to Fig. 6, the neighboring pixels of pixel P, pixels P1, P2, P3, P4, P5, P6, P7 and P8 are selected, so the surrounding pixels of pixel P are used. In the present embodiment, the gray level of each pixel in the image is from 0 to 255. Wherein the gray level of pixel P is 122, while the gray level of pixels P1, P2, P3, P4, P5, P6, P7 and P8 are 136, 124, 119, 130, 125, 132, 124, and 120 respectively.

According to Sobel operator, with the gray levels of the neighbor pixels P1, P2, ... P8, etc., the gradient of gray level of pixel P is given by formula (1):

$$DX1=(G3+2\times G5+G8)-(G1+2\times G4+G6);$$
$$DY1=(G1+2\times G2+G3)-(G6+2\times G7+G8). \qquad (1)$$

Where DX1 is the x component of the gradient of the gray level distribution at pixel P, DY1 is the y component of the gradient of the gray level distribution at P; G1, G2, ... G8 are the gray levels of pixels P1, P2, ... P8 respectively. Thus, on the basis of Formula (1), the gradient of gray level of pixel P can be determined as (-39, 3).

Similarly, the gradients of the gray levels of other pixels in the annular region 300R can be calculated. In the present embodiment, the gradients of pixels P1, P2, P3, P4, P5, P6, P7 and P8 are determined as (-30, 5), (-36, 4), (-32, 1), (-33,6), (-30, 4), (-38, 8), (-33, 3) and (-31, -2) respectively. Thus, the gradient of gray level of each pixel in the annular region 300R can be determined in the same way.

Returning to Fig. 3, in step S43, the first determination means 212 also determines the weight of the gradient of gray level of each pixel in the annular region 300R. For example, the gradient of gray level at a pixel can be given by formula (2):

$$W1=(|DX1| + |DY1|)/255; \qquad (2)$$

Where W1 denotes the weight of a pixel, DX1 is the x component of the gradient of the gray level at the pixel, and DY1 is the y component of the gradient of the gray level at the pixel.

For example, referred to Fig. 6, in the present embodiment, the gradient of the gray level of pixel P is known as (-39, 3), hence the weight of the gradient of gray level of pixel P is $(|-39| + |3|)/255=0.165$.

Likewise, the weights of the gradient of gray level of other pixels P1,

P2, P3, P4, P5, P6, P7, P8, etc., in the annular region 300R can be easily determined as 0.137, 0.157, 0.129, 0.153, 0.133, 0.180, 0.141, 0.129, etc. respectively.

Determining the Ellipse Gradient of Each Pixel in the Annular region

Returning to Fig. 4, the flow goes to step S44 after step S43. In step S44, the second determination means 213 determines the ellipse gradient of each pixel in the above-mentioned annular region 300R, which is to be used in determining the degree how the portion of image in annular region 300R is closed to a human face image.

Next, referring to Fig. 7, how to determine the ellipse gradient of each pixel in annular region 300R by the second determination means 213 is explained below.

Fig. 7 is a diagram showing the ellipse gradient of gray level of each pixel in the annular region 300R.

Referring to Fig. 7, the coordinate of pixel P in the given coordinate system is known as (380, 250), and reference numeral E11 denotes an inscribed ellipse of rectangular region 300A, and reference numeral E12 denotes a second ellipse centered at the center of rectangle 300A and passing pixel P. As to the ellipse gradient of a certain pixel P in the annular region 300R, the second ellipse E12 can be determined according to the inscribed ellipse E11 in the given rectangular region 300A. Wherein the center of the determined second ellipse E12 is the same as that of the inscribed ellipse E11, the ratio of major axis to minor axis is equal; and the second ellipse E12 passes through said pixel P. Wherein the ellipse gradient of the pixel is given

by formula (3):

$$Dex=(X_P-X_A)/(W/2)^2;$$
$$Dey=(Y_P-Y_A)/(L/2)^2 \qquad (3)$$

Wherein Dex is the x component of the ellipse gradient of pixel P, Dey is the y component of the ellipse gradient of the pixel, X_P denotes the X coordinate of said pixel, X_A denotes the X coordinate of the center of the rectangular region, W denotes the width of the rectangular region; Y_P denotes the Y coordinate of said pixel, Y_A denotes the Y coordinate of the center of the rectangular region, and L denotes the length of the rectangular region.

In the present example, the coordinate of the pixel P in the given coordinate system is (380, 250). Further, the width and length of the rectangular region 300A in the present example is 140 and 176 respectively. The coordinates of the center of the rectangular region 300A is (300, 300). Therefore, the ellipse gradient of the pixel P is determined as (0.016, -0.006) according to the above-mentioned formula (3).

Likewise, the ellipse gradient of other pixels P1, P2, P3, P4, P5, P6, P7, P8, etc. in the annular region 300R can be determined as (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), etc. respectively.

After having acquired the ellipse gradient of each pixel in the annular region 300R, that is, after step S44, the flow goes to step S45. In step S45, the angle between the gradient of gray level and its corresponding ellipse gradient of each pixel in the annular region 300R is determined. In the present example, the gradient of gray level of pixel P is given to be (-39, 3), and its

ellipse gradient to be (0.016, -0.006), the angle θ of the two gradients of pixel P is determined by formula (4):

$$\theta=\cos^{-1}\{(DX1 \times Dex + DY1 \times Dey)/[(DX1^2+DY1^2) \cdot (Dex^2+Dey^2)]^{1/2}\} \quad (4)$$

Thus, the angle between the two gradients of pixel P is determined as 0.29 (radian). Similarly, the angles between the two gradients at other pixels90

P1, P2, P3, P4, P5, P6, P7, P8, etc. in the annular region 300R can be easily determined as 0.19, 0.26, 0.33, 0.18, 0.23, 0.50, 0.27, 0.42 (radian), etc. respectively.

Further, in step S45, the average of the angles between the gradient of gray level and its corresponding ellipse gradient for all pixels in the annular region 300R is given by formula (5):

$$A1=S1/C1 \quad (5)$$

Where A1 denotes the average of the angles between the gradient of gray level and its corresponding ellipse gradient for all pixels in the annular region, S1 denotes the sum of the angles between the two gradients of all pixels in the annular region; C1 denotes the total number of pixels in the annular region.

As to the present example, the average angle between the gradient of gray level and its corresponding ellipse gradient of all pixels in the annular region 300R is determined as 0.59.

Returning to step S45 in Fig. 3, after the above-mention average of the gradient differences has been determined, it is determined in step S45 whether said average is less than a first threshold. In the present invention, the first threshold is between 0.01 and 1.5, such as 0.61. If the calculated average is determined to be less than the first threshold, then the flow goes to step S46. If not, the flow goes to step S48. In step S48, it is determined that the rectangle 300A does not contain an image of a human face, and then, the flow ends in step S49.

In the present example, the average of the angle between the gradient of gray level and its corresponding ellipse gradient for all the pixels in the annular region 300R is 0.59, which is less than the first threshold. Therefore, the flow goes to step S46.

In step S46, using the weight W1 as described above, a weighted average of the angle difference and the gradient of gray level and the ellipse gradient for all the pixels in the annular region 300R is determined, as is given by formula (6):

$$A2=S2/C2; \qquad (6)$$

Where A2 denotes the weighted average of the angle between the two gradients for all the pixels in the annular region, C2 denotes the total number of pixels in the annular region, S2 denotes the sum of the product of the angle and the weight of gradient of gray level for all the pixels in the annular region.

In the present example, the weighted average angle between the two gradients for all pixels in the annular region is 0.66.

After the determination of the weighted average angle, it is determined in step S46 whether said weighted average angle is less than a second threshold, such as 0.68. The second threshold in the present invention is ranged from 0.01 to 1. If the weight average angle is determined to be less than the second threshold, then the flow goes to step S47, where it is determined that the rectangle 300A contains an image of a human face, then the flow ends in step S49. If the weighted average angle is not less than the second threshold, then the flow goes to step S48. In step S48, it is determined that the rectangle 300A does not contain an image of a human face, and then, the flow ends in step S49.

As to the present example, the weighted average angle for all pixels in the annular region 300R is 0.66, which is less than the second threshold, and then the flow goes to step S47. In step S47, and the region 300A is determined to contain an image of a human face. Afterwards, the flow ends in step S49.

The Second Example

Fig. 8A is a diagram showing another example of an original image to be detected. The original image 800 in Fig. 8A is produced by a digital camera. Of course, it also can be input to human face identification system by a digital device 111, such as a scanner or the like. Assume that the original image is stored in a predetermined region in the HD 107 or the RAM 109, or the like.

Fig. 8B is a diagram showing a rectangular region 800B in the original image 800, which is shown in Fig. 8A, and the location in Cartesian

coordinate system, whereat the X and Y axes extend in horizontal and vertical direction respectively.

As to the rectangular region 800B shown in Fig. 8B, the same flow chart shown in Fig. 4 is adopted to explain the process of determining whether the rectangular region 800B contains a human face.

Referring to Fig. 4, eye detecting device 218 determines eyes in the image, in step S41, reading means 210 reads the gray level of each pixel in the original image 800, and the sub-image determining device 219 determines rectangular regions to be identified, such as rectangular region 800B, in the original image 800. As shown in Fig. 8A, the origin of the Cartesian coordinate system is taken as the upper left corner of the original image 800. Then the coordinate of each of the four corners of rectangular region 800B shown in Fig. 8B is (203, 219), (373, 219), (203, 423) and (373, 423) respectively in the Cartesian coordinate system. That is, its width and length is 170 and 204 respectively.

Then, in step S42, the annular region 800C around the rectangular region 800B in Fig. 8B is determined by the annular region determination means 211. Referred to Fig. 8C, as to the rectangular region 800B by scaling, its width and length with a factor of 9/7 and a factor of 5/7 respectively, two rectangular regions 800C1 and 800C2 are generated.

Fig. 8C is a diagram showing the annular region 800C, which is determined for the rectangular region 800B. The annular region 800C is formed by the first rectangular region 800C1 and the second rectangular region 800C2. The coordinates of each of the four corners of first rectangular region 800C1 in this coordinate system are (170, 190), (397, 190), (179, 452)

and (397, 452) respectively, while the coordinates of each of the four corners of the second rectangular region 800C2 in this coordinate system are (227, 248), (349, 248), (227, 394) and (349, 394) respectively.

Then, in step S43, the gradient of the gray level of each pixel in the annular region 800C is determined by the first determination means 212, which is (188, 6), (183, 8), (186, 10), (180, 6), (180, 6), etc. respectively. Similarly, the weight of the gradient of the gray level of each pixel in the annular region 800C is determined in the same way, which is 0.76, 0.75, 0.77, 0.77, 0.73, etc. respectively.

Further, in step S44, the ellipse gradient of each pixel in the annular region 800C is determined respectively by a second determination means 213, which (-0.015, -0.012), (-0.015, -0.012), (-0.015, -0.012), (-0.014, -0.012), etc. respectively. Then, the flow goes to step S45.

In step S45, the angle between the two gradients of each pixel in the annular region 800C is determined; the average angle between the two gradients for pixels in the annular region 800C can be determined in the same way as above-mentioned, which is 0.56 for the present example and is less than the first threshold. Then the flow goes to step S46.

In step S46, the weighted average angle between the two gradients for pixels in the annular region 800C is determined as 0.64, which is less than the first threshold, and the flow goes to step S47, where it is determined that the rectangular region 800B contains an image of a human face. Then, the flow ends in step S49.

The Third Example

Fig. 8D is a diagram showing a rectangular region 800D is the original image 800, which is shown in Fig. 8A.

As to the rectangular region 800D shown in Fig. 8D, the flow chart shown in Fig. 4 is used to illustrate the process of determining whether the rectangular region 800D contains an image of a human face.

Referring to Fig. 4, firstly in step S41, reading means 210 reads the gray level of each pixel in an original image 800; eye detecting means 218 determines eyes in the image; and the sub-image determining device 219 determines the rectangular regions to be identified in the image, such as rectangle 800D in the original image 800. As shown in Fig. 8A, the origin of the Cartesian coordinate system is taken as the upper left corner of the original image 800. The coordinates of four corners of the rectangular region 800D shown in Fig. 8D are (439, 158), (533, 158), (439, 256) and (533, 256) respectively, namely its width and length are 94 and 98 respectively.

Then in step S42, the annular region 800E in the rectangular region 800D in Fig. 8D is determined by the annular region determination means 211. Referring to Fig. 8E, as to the rectangular region 800D by scaling, its width and length with a factor of 9/7 and a factor of 5/7 respectively, two rectangular regions 800E1 and 800E2 are determined.

Fig. 8E is a diagram showing the annular region 800E, which is determined for the rectangular region 800D. The annular region 800E is formed by the first rectangular region 800E1 and the second rectangular region 800E2. The coordinates of each of the four corners of the first

rectangular region 800E1 in this coordinate system are (426, 144), (546, 144), (426, 270) and (546, 270) respectively, while the coordinates of each of the four corners of the second rectangular region 800E2 in this coordinate system are (452, 172), (520, 172), (452, 242) and (520, 242) respectively.

Then, in step S43, the gradient of the gray level of each pixel in the annular region 800E is determined by a first determination means 212, which is (241, 27), (246, 22), (248, 16), (246, 14), (253, 15), etc. respectively; while the weight of the gradient of the gray level of each pixel in the annular region 800E is determined in the same way, which is 1.05, 1.05, 1.05, 1.04, 1.05, etc. respectively.

Then, the flow goes to step S44. In step S44, the ellipse gradient of each pixel in the annular region 800E is determined respectively by a second determination means 213, which is determined as (-0.027, -0.027), (-0.027, -0.027), (-0.026, -0.027), (-0.026, -0.027), (-0.025, -0.027), etc. respectively, then the flow goes to step S45.

In step S45, the angle between the gradient of the gray level distribution of the image and the ellipse gradient as determined above for each pixel in the annular region 800E is determined, which is 0.71, 0.73, 0.74, 0.76, 0.77, etc. respectively; the average angle between the two gradients for each pixel in the annular region 800E is calculated in the same way as in the above example, which is 0.64. In the present example, since the average is 0.64, which is not less than the first threshold, the flow goes to step S48, where it is determined that the rectangular region 800D does not contained an image of a human face. Then, the flow ends in step S49.

Alternative Embodiment

An alternative embodiment of the present invention is illustrated below.

Fig. 9 is a diagram showing the human face determination process according to the alternative embodiment of the present invention. In Fig. 9, the same reference numeral as that in Fig. 4 denotes the same process.

Again, the original image 800 in Fig. 8A is taken as an example to illustrate determining whether a rectangular region 800B in the original image 800 contains a human face image.

First, as the same as that in Fig. 4 the process flow starts in step S40. In step S41, reading means 210 reads the gray level of each pixel in original image 800; eye detecting means 218 detect eye areas in the image; and sub-image determining means 219 determines the sub-image(s) based on the eye areas detected, such as rectangular region 800B, to be identified. Then the flow goes to step S42. In step S42, a annular region 800C is determined for rectangular region 800B, as shown in Fig. 8C. Further, the flow goes to step S43. In step S43, the gradient of the gray level at each pixel in the annular region 800C is determined, which is (188, 6), (183, 8), (186, 10), (180, 6), (180, 6), etc. respectively. The weight of the gradient of the gray level at each pixel in the annular region 800C is determined in the same way as in the previous embodiment, which is 0.76, 0.75, 0.77, 0.77, 0.73, etc. respectively. Afterwards, the flow goes to step S44. In step S44, the ellipse gradient of each pixel in the annular region 800C is determined, which is (-0.015, -0.012), (-0.015, -0.012), (-0.015, -0.012), (-0.015, -0.012), (-0.014, -0.012), etc. respectively.

Then, the flow goes to step S95. In step S95, the angle between the

gradient of the gray level and its corresponding ellipse gradient at each pixel in the annular region 800C is calculated respectively e.g., in the same way as the previous embodiment. Which is 0.64, 0.63, 0.62, 0.64, 0.64, etc. (in radian), respectively. Then, the weighted average angle between the two gradients for all pixels in the annular region 800C is calculated. While in step S95, it is determined whether the weighted average is less than a third threshold, such as 0.68. In the present invention, the third threshold is ranged from 0.01 to 1. If the weighted average angle is less than the third threshold, then the flow goes to step S96. In step S96, the rectangular region is determined to contain a human face image. If the weighted average angle is not less than the third threshold, then the flow goes to step S97, where it is determined that region 800D does not contain a human face image. Then the flow ends in step S98.

In the present embodiment, since the mean of the product of the gradient difference and its corresponding gradient weight of each pixel in the annular region is 0.64, which is less than the third threshold, the flow goes to step S96 to determine that the rectangular region to be determined 800B contains a human face. Afterwards, the flow ends in step S98.

In the method and apparatus for human face identification, human face identification is performed by evaluating the difference between the direction of the gradient of the gray level distribution of the image to be processed and the direction of a reference gradient, such as the ellipse gradient as described above. Thus, the magnitudes of the gradient of the gray level distribution and the reference gradient do not affect the results of the evaluation.

Thus, other reference gradient may be determined for a sub-image 300A for human face identification.

For example, for each sub-image 300A determined, a reference gray-level distribution of a human face, which can be expressed as:

$$z(x,y)= -(x^2/a^2+y^2/b^2)+h$$

is determined, wherein h is a constant, a/b equals to the ratio of the width of said sub-image to the height of said sub-image, the origin is at the center of said sub-image, and z(x,y) is the gray-level value at point P(x,y). Figure 10 shows image with such a reference gray-level distribution.

Then, an annular region 300R is generated, as explained above. After that, gradients ∇z and ∇g are calculated. ∇g is the gradient of the gray level distribution and can be calculated using any operator for discrete gradient calculation, e.g. Sobel operator. Then, the angle between vectors ∇z and ∇g is calculated.

Specifically, the angle θ between ∇z and ∇g at point P(x,y) can be calculated using:

$$\cos\theta = \nabla z \cdot \nabla g / (|\nabla g| \times |\nabla z|) \text{ and}$$
$$\theta = \cos^{-1}(\nabla z \cdot \nabla g / (|\nabla g| \times |\nabla z|))$$

so that the correspondence between the two vector fields ∇z and ∇g can be evaluated as explained above.

As a further alternative embodiment the present invention, for point P(x,y), the direction of ∇g and the tangent direction of the curve

$$-(x^2/a^2+y^2/b^2)+h=p$$

which passes through point P(x,y) and which is the contour line of z(x,y)=p, (where p is a parameter), and then evaluate the correspondence between the direction of ∇g and the tangent (or normal) direction of the curve $-(x^2/a^2+y^2/b^2)+h=p$, e.g. by calculating the angle between the two directions, while p varies to generate a group of ellipses so as to cover the entire annular area 300R, and then determine whether the image contains human face(s) based on the result of the evaluation.

Note that the present invention may be applied to either a system constituted by a plurality of devices (e.g., a host computer, an interface device, a reader, a printer, and the like), or an apparatus consisting of a single equipment (e.g., a copying machine, a facsimile apparatus, or the like).

The objects of the present invention are also achieved by supplying a storage medium. The storage medium records a program code of a software program that can implement the functions of the above embodiment to the system or apparatus, and reading out and executing the program code stored in the storage medium by a computer (or a CPU or MPU) of the system or apparatus. In this case, the program code itself read out from the storage medium implements the functions of the above-mentioned embodiment, and the storage medium, which stores the program code, constitutes the present invention.

As the storage medium for supplying the program code, for example, a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM, CD-R, magnetic tape, nonvolatile memory card, ROM, and the like may be used.

The functions of the above-mentioned embodiment may be implemented not only by executing the readout program code by the computer but also by some or all of actual processing operations executed by

an OS (operating system) running on the computer on the basis of an instruction of the program code.

As can be seen from the above, the method of the present invention provides a fast approach for determining human face in a picture with a complex background, without requiring the detected picture to have a very high quality, thereby substantially eliminating the possibility of the human face in the picture being skipped over. The method allows for the precision determination of human face under different scales, orientation and lighting condition. Therefore, in accordance with the present invention, with the method, apparatus or system, the human face in a picture can be quickly and effectively determined.

The present invention includes a case where, after the program codes read from the storage medium are written in a function expansion card which is inserted into the computer or in a memory provided in a function expansion unit which is connected to the computer, CPU or the like contained in the function expansion card or unit performs a part or entire process in accordance with designations of the program codes and realizes functions of the above embodiment.

In a case where the present invention is applied to the aforesaid storage medium, the storage medium stores programs codes corresponding to the flowcharts (Fig. 4, and 9) described in the embodiments.

The embodiment explained above is specialized to determine human face, however, the present invention is not limited to determine human face, it is applicable to other determination method, for example, method to detect flaw portion on a circuit board.

As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.

Claims

1. A method for determining an object in an image having a gray-level distribution,

characterized in that said method comprises the steps of:

a)for a subset of pixels in said image, deriving a first variable from said gray-level distribution of said image,

b)for said subset of pixels, deriving a second variable from a preset reference gray-level distribution, said reference gray-level distribution being characteristic of said object;

c)evaluating the correspondence between said first variable and said second variable over the subset of pixels;

d)determining if said image contains said object based on the result of the evaluation step.

2. The method of claim 1, characterized in that said step of selecting a subset of the pixels in said image further comprises the steps of:

selecting an area in said image,

wherein pixels of said subset are in said area, and

wherein over pixels in said area the first variable as derived from the gray-level distribution of the object to be detected has better correspondence to the second variable than over pixels outside said area.

3.The method of claim 2, characterized in that it further comprises the step of:

determining a sub-image in said image; and

selecting said area in said sub-image.

4.The method of any one of claim 1-3, wherein said first variable

represents the direction of the gradient of the gray-level distribution of said image.

5.The method of any one of claim 4, wherein said second variable represents the direction of the gradient of said reference gray-level distribution.

6.The method of claim 5, wherein said gradient is a discrete gradient.

7.The method of any one of claim 5, characterized in that said evaluation step includes a statistical processing over said subset of pixels.

8.The method of claim 5, characterized in that said evaluation step includes a statistical processing over said subset of pixels and said statistical process comprises a weighting process, in which pixels with larger magnitude of gradient of the first variable have larger weights.

9.The method of claim 8, characterized in that said evaluating step includes both the statistical process comprising the weighting process and another statistical process without the weighting process and determines whether the image contains the object in accordance with the results of both of the statistical processes.

10.The method of claim 3, characterized in that said area is an annular area, which centers around the center of said sub-image.

11.The method of claim 2, wherein said subset of pixel includes all the pixels in said area.

12.The method of claim 3, characterized in that said sub-image is determined by identifying a predetermined feature in said image.

13.The image of claim 12, characterized in that said predetermined feature includes a pair of dark areas.

14.The method of claim 13, characterized in that said sub-image is determined based on the positions of and interval between the centers of said pair of dark areas.

15.The method of claim 13, characterized in that said reference gray-level distribution is expressed by

$$z(x,y)= -(x^2/a^2+y^2/b^2)+h$$

wherein h is a constant, a/b equals to the ratio of the width of said sub-image to the height of said sub-image, and the origin is at the center of said sub-image.

16. The method of claim 5 or 15, wherein said evaluating step includes:

for all the pixels in said subset, calculating the average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference gray-level distribution;

comparing said average value with a predetermined value; and

determining that the sub-image contains said object if said average value equals to or is smaller than said predetermined value.

17. The method of claim 5 or 15, wherein said evaluating step includes:

for all the pixels in said subset, calculating a weighted average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference gray-level distribution, wherein a pixel having a greater magnitude of gradient of gray-level distribution has a larger weight;

comparing said weighted average value with a predetermined value; and

determining that the sub-image contains said object if said weighted average value equals to or is smaller than said predetermined value.

18. The method of claim 5 or 15, wherein said evaluating step includes:

for all the pixels in said subset, calculating the average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference gray-level distribution;

for all the pixels in said subset, calculating a weighted average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference gray-level distribution, wherein a pixel having a greater magnitude of gradient of gray-level distribution has a larger weight;

comparing said average value with a first predetermined value;

comparing said weighted average value with a second predetermined value; and

determining that the sub-image contains said object if said average value equals to or is smaller than said first predetermined value and said weighted average value equals to or is smaller than said second predetermined value.

19. A method for determining an object in an image having a gray-level distribution, characterized in that said method comprises the steps of:

a) determining a sub-image in said image;

b) for a subset of pixels, deriving a first variable from said gray-level distribution of said image,

c) for said subset of pixels, deriving a second variable from a preset

reference gray-level distribution, said reference gray-level distribution being characteristic of said object;

d) evaluating the correspondence between said first variable and said second variable over the subset of pixels; and

e) determining if said image contains said object based on the result of the evaluation step.

20. A storage medium with a program code for object-identification in an image with a gray-level distribution stored therein, characterized in that said program code comprises:

codes for determining a sub-image in said image;

codes for deriving, for a subset of pixels, a first variable from said gray-level distribution of said image,

codes for deriving, for said subset of pixels, a second variable from a preset reference gray-level distribution, said reference gray-level distribution being characteristic of said object;

codes for evaluating the correspondence between said first variable and said second variable over the subset of pixels; and

codes for determining if said image contains said object based on the result of the evaluation step.

21. An image processing method for determining a feature portion in an image, comprising:

a read step of reading the image and a rectangle region to be determined in the image;

a setting step of setting a band region surrounding said rectangle region;

a first determination step of determining the gradient of gray level of each pixel in said band region;

a second determination step of determining an ellipse for each pixel in said

band region, said ellipse passing through a pixel in said band region, and determining the ellipse gradient of each pixel in the band region; and

a third determination step of determining the feature portion contained in said rectangle region on the basis of the gradient of the gray level of each pixel and the ellipse gradient of each pixel in said band region.

22. The image processing method according to claim 21, wherein the image contained said rectangle region to be determined is a feature portion.

23. The image processing method according to claim 21, wherein said band region is formed by a first rectangle region and a second rectangle region.

24. The image processing method according to claim 23, wherein said first rectangle region is located inside said rectangle region to be determined.

25. The image processing method according to claim 23, wherein said second rectangle region is located outside said rectangle region to be determined.

26. The image processing method according to claim 23, wherein the center of said first rectangle region lies in the same location as that of said rectangle region to be determined, and the width and the length of said first rectangle region respectively equals m times the width and the length of said rectangle region to be determined, where $0 < m < 1$.

27. The image processing method according to claim 23, wherein the center of said second rectangle region lies in the same location as that of said rectangle region to be determined, and the width and the length of said second rectangle region respectively equals n times the width and the length of said rectangle region to be determined, where $1 < n < 2$.

28. The image processing method according to claim 21, wherein said first determination step comprising the step of determining the gradient of the gray level of a pixel on the basis of the gray level of the pixels surrounding the pixel.

29. The image processing method according to claim 28, wherein the number of pixels surrounding said pixel equals k^2, where k is an integer, and $2 < k < 15$.

30. The image processing method according to claim 21, wherein the center

of said ellipse lies in the same location as that of the inscribed ellipse in the rectangle region to be determined, and the ratio of the major axis to the minor axis of said ellipse is the same as that of the inscribed ellipse in the rectangle region to be determined.

31. The image processing method according to claim 21 or 30, wherein further comprising the step of determining the difference of the gradient of the gray level of each pixel and the ellipse gradient of each pixel in said band region.

32. The image processing method according to claim 21 or 30, wherein further comprising the step of determining the mean of the difference of the gradient of the gray level of each pixel and the ellipse gradient of each pixel in said band region.

33. The image processing method according to claim 21 or 32, wherein further comprising the step of determining the weight of the gradient of the gray level of each pixel in said band region.

34. The image processing method according to claim 32, wherein further comprising the step of determining whether the mean of the difference of the gradient of the gray level of each pixel and the ellipse gradient of each pixel in said band region is less than the first threshold, where the first threshold is within the range of 0.01 to 1.5, preferably of 0.61.

35. The image processing method according to claim 34, wherein further comprising the step of determining whether the mean of the product of the difference, which is that of the gradient of the gray level of each pixel and the ellipse gradient of each pixel in said band region, and the weight of the gradient of the gray level of each pixel in said band region is less than the second threshold, where the second threshold is within the range of 0.01 to 1, preferably of 0.68.

36. The image processing method according to claim 33, wherein further comprising the step of determining whether the mean of the product of the difference, which is that of the gradient of the gray level of each pixel and the gradient of its said corresponding ellipse, and the weight of the gradient of

the gray level of each pixel in said band region is less than the third threshold, where the third threshold is within the range of 0.1 to 1, preferably of 0.68.

37.A human eye detection method for detecting human eye in an image, comprising:

the read step of reading the gray level of each pixel in the each column in the image;

the segment step of segmenting each column into a plurality of intervals , and labeling each of the intervals as valley region , relay region or peak region;

the merging step of merging the valley region of the each column and the valley region of its adjacent column, and generating an eye candidate region; and

the determination step of determining the human eye from the eye candidate regions.

38.The method according to claim 37, wherein the segment step includes the step of labeling each interval as one of valley region, relay region and peak region based on the gray level of each the interval in a column.

39.The method according to claim 37, wherein the segment step includes the step of labeling each interval as one of valley region, relay region and peak region based on the luminance value of each interval in a column.

40.The method according to claim 37, wherein the segment step includes the step of labeling each interval in a column as one of valley region, relay region and peak region based on the respective gray level ratios of the interval to its two adjacent intervals.

41.The method according to claim 37, wherein the segment step includes the step of labeling each interval in a column as one of valley region, relay region and peak region based on the respective luminance value ratios of the interval to its two adjacent interval.

42.The method according to any one of claims 37, wherein further

comprising the step of comparing the gray level of a segment point with that of its two adjacent intervals, and adding the segment point into the adjacent interval which has a gray level close to that of the segment point.

43.The method according to any one of claims 37, wherein further comprising the step of comparing the gray level of a relay region with that of its adjacent valley region and peak valley, and adding the relay region into the valley region on the basis of the gray levels of the adjacent valley region, peak region and the relay region.

44.The method according to claim 37, wherein the merging step comprising the steps of

setting each valley region in the first column of the image as a seed region respectively;

reading out the valley region of next column of the image,

determining whether the valley region can be merged into the seed region,

merging the valley region that can be merged into the seed region;

setting the non-merged valley region as a seed region; and

determining the seed region which has no further merged valley region as an eye candidate region.

45.The method according to claim 44, wherein further comprising the step of comparing the size of overlap of the valley region and the seed region's predicted valley region, the valley region and the seed region's predicted valley region,

46.The method according to claim 44, wherein further comprising the step of comparing the gray level of the valley region and that of the seed region.

47.The method according to claim 44, wherein further comprising the step of comparing the gray levels of the valley region, the seed region, the background gray levels of the valley region and the seed region.

48.The method according to claim 37, wherein the determination step comprising the step of determining whether the gray level of the eye candidate region is less than the first threshold, where the first threshold is within the range of

100 to 200, preferably of 160.

49.The method according to claim 48, wherein the determination step further comprising the step of determining whether the background gray level of the eye candidate region is bigger than the second threshold, where the second threshold is within the range of 20 to 80, preferably of 30.

50.The method according to claim 49, wherein the determination step further comprising the step of determining whether the difference of background gray level of the eye candidate region and its gray level is bigger than a third threshold, where the third threshold is within the range of 5 to 120 , preferably of 20.

51. The method according to claim 50, wherein the determination step further comprising the step of determining whether the ratio of the width to the height of an eye candidate region is bigger than a fourth threshold, where the fourth threshold is within the range of 1 to 5, preferably of 3.33.

52.The method according to claim 51, wherein the determination step further comprising the step of determining whether the ratio of the size of an eye candidate region to that of its circum-rectangle is bigger than the fifth threshold, where the fifth threshold is within the range of 0.2 to 1 preferably of 0.4.

53. The method according to claim 14, wherein the step for determining a sub-image in said image includes:

reading the gray level of each pixel in the each column in the image;

segmenting each column into a plurality of intervals , and labeling each of the intervals as valley region , relay region or peak region;

merging the valley region of the each column and the valley region of its adjacent column, and generating a candidate dark area; and

determining a dark area.

54. The method according to claim 53, wherein the merging step comprising the steps of

setting each valley region in the first column of the image as a seed region;

reading out the valley region of the next column of the image;

determining whether the valley region can be merged into the seed region,

merging the valley region that can be merged into the seed region;

setting the valley region that cannot be merged into the seed region as a seed region; and

determining the seed region which has no further merged valley region as a dark area.

ABSTRACT OF THE INVENTION

A human face determination method, apparatus, and system for exactly determining the human face in an image are disclosed in the present invention. The said face determination apparatus comprises a read means; a setting means; a first determination means; a second determination means; and a third determination means. According to the present invention, a reliable and fast method for determining human face with a complex background in an image is implemented.

determining device 21~~8~~9 determines one or more sub-images (or regions) 300A in the original image 300 for human face identification, and it also determines the location of the sub-image 300A in the original image 300. The sub-image 300A can be substantially rectangular and is the candidate region of a human face image portion. However, the shape of the sub-image is not limited to rectangular but can be any other suitable shape.

A method and device, in accordance with the present invention, for determining sub-image(s) 300A in an image for identification are described below. It is to be noted, however, that the manner of eye area detection for determining the sub-image is not limited to the method of the present invention as described below. Instead, other methods and/or processes, as those known in the art, can be utilized for determining the sub-image.

Determination of Sub-image 300A

Eye Detection Device

Fig. 12 is a block diagram showing the arrangement of the eye-detecting device according to the an embodiment of the present invention;

The eye detection device 1114 of the present embodiment comprises ~~a reading means 1200,~~ a segment means 1201, a merger means 1202 and a determination means 1203. ~~In the eye detection device 1114, a the reading means 1200 performs an image-reading process, in which the gray level of each pixel in the columns of an image stored in, for example, HD 1107 is read by the reading means 1200.~~ Referring to Figs. 14D and 14E, on the basis of the gray level of each pixel in a column C41 of an image, the column C41 of an image is segmented into a plurality of intervals I1-1, I1-2, ... I1-9, I1-10 by segment means 1201. These intervals I1-1, I1-2, ... I1-9, I1-10 can be marked as classified into three types: peak regions, valley regions and

intermediate regions, according to their average gray level of the image data. The terms "peak region, valley region and intermediate region" will be defined in detail later. Then, the valley regions of column C41 can be obtained. In the same way, the segment means 1201 also divides other columns of the image into the three types and obtains their valley regions respectively. After all the columns of an image have been marked as the three types and their valley regions have been obtained, the merger means 1202 executes the merging process and merges the valley regions in the adjacent columns. The merged valley regions are set as the human eye candidates. Then the human eye can be determined by determination means 1203.

Detecting the eye areas

A human eye detection process for an original image will be explained below with reference to the flow chart in Fig. 13A. Fig. 13B is an example of an original image to be detected. Assume that the original image is stored in a predetermined area in the HD 1107 or the RAM 1109, or the like.

Referring to Fig. 13A, in ~~step S131, the reading means 1200 reads the gray level of each pixel in the columns of the original image to be detected. If the original image is encoded by, e.g., JPEG, the reading means 1200 must first decode it before reads its image data. Of course, all the gray levels for every columns of the original image can be read at one time, or they can be read sequentially or respectively. In the preferred embodiment, each column of the image is read, . However, each row of the image can be read, as in case the given image has been related to 90°. In~~ step S132, every columns of the original image are is segmented into many intervals by the segment means 1201. With reference to Fig. 14E, the lengths of each of the intervals I1-1, I1-2, ... I1-9 and I1-10 is variable. For example, the length of interval I1-1 is not equal to the length of interval I1-2. Some of the segmented intervals are marked as the valley regions on the basis of their average gray levels of pixels.

P2, P3, P4, P5, P6, P7, P8, etc., in the annular region 300R can be easily determined as 0.137, 0.157, 0.129, 0.153, 0.133, 0.180, 0.141, 0.129, etc. respectively.

Determining the ~~Ellipse~~Reference Gradient of Each Pixel in the Annular region

Returning to Fig. 4, the flow goes to step S44 after step S43. In step S44, the second determination means 213 determines ~~the ellipse~~a reference ~~gradient~~ distribution of each pixel in the above-mentioned annular region 300R in accordance to the sub-image 300A~~.~~; the reference distribution~~ which~~ is to be used in determining~~ the degree~~ how the portion of image in the annular region 300R is closed to a human face image.

In the method and apparatus for human face identification of the present invention, human face identification is performed by evaluating the difference between the direction of the gradient of the gray level distribution of the image to be processed and the direction of a reference gradient derived from the reference distribution.

Specifically, for each sub-image 300A determined, a reference distribution of a human face is determined, whic can be expressed as:

$$z(x,y) = -\frac{(x-x_c)^2}{a^2} - \frac{(y-y_c)^2}{b^2} + h$$

wherein h is a constant, a/b equals to the ratio of the width of said sub-image to the height of said sub-image, (x_c, y_c) is the center of the sub-image. Figure 10 shows an image with such a reference distribution, and Figure 7 shows the contours of (E11 and E12) of the reference distribution.

Then, gradients ∇z and ∇g are calculated. ∇g is the gradient of the gray

level distribution and can be calculated using any operator for discrete gradient calculation, e.g. Sobel operator. Then, the angle between vectors ∇z and ∇g is calculated.

Specifically, the angle θ between ∇z and ∇g at point (x,y) can be calculated using:

$$\cos\theta = \nabla z \cdot \nabla g / (|\nabla g| \times |\nabla z|) \qquad (7)$$

$$\theta = \cos^{-1}(\nabla z \cdot \nabla g / (|\nabla g| \times |\nabla z|)) \qquad (8)$$

~~Next, referring to Fig. 7, how to determine the ellipse gradient of each pixel in annular region 300R by the second determination means 213 is explained below.~~

~~Fig. 7 is a diagram showing the ellipse gradient of gray level of each pixel in the annular region 300R.~~

~~Referring to Fig. 7, the coordinate of pixel P in the given coordinate system is known as (380, 250), reference numeral E11 denotes an inscribed ellipse of rectangular region 300A, and reference numeral E12 denotes a second ellipse centered at the center of rectangle 300A and passing pixel P. As to the ellipse gradient of a certain pixel P in the annular region 300R, the second ellipse E12 can be determined according to the inscribed ellipse E11 in the given rectangular region 300A. The center of the determined second ellipse E12 is the same as that of the inscribed ellipse E11, and the two ellipses have the same ratio of major axis to minor axis; and the second ellipse E12 passes through said pixel P. The ellipse gradient of the pixel is~~

~~given by formula (7):~~

~~$Dex=(X_P-X_A)/(W/2)^2$;~~

~~$Dey=(Y_P-Y_A)/(L/2)^2$ —————————— (7)~~

~~Wherein Dex is the x component of the ellipse gradient of pixel P, Dey is the y component of the ellipse gradient of the pixel, X_P denotes the X coordinate of said pixel, X_A denotes the X coordinate of the center of the rectangular region, W denotes the width of the rectangular region; Y_P denotes the Y coordinate of said pixel, Y_A denotes the Y coordinate of the center of the rectangular region, and L denotes the length of the rectangular region.~~

~~In the present example, the coordinate of the pixel P in the given coordinate system is (380, 250). Further, the width and length of the rectangular region 300A in the present example is 140 and 176 respectively. The coordinates of the center of the rectangular region 300A is (300, 300). Therefore, the ellipse gradient of the pixel P is determined as (0.016, -0.006) according to the above-mentioned formula (7).~~

~~Likewise, the ellipse gradient of other pixels P1, P2, P3, P4, P5, P6, P7, P8, etc. in the annular region 300R can be determined as (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), (0.016, -0.006), etc. respectively.~~

~~After having acquired the ellipse gradient of each pixel in the annular region 300R, that is, after step S44, the flow goes to step S45. In step S45, the angle between the gradient of gray level and its corresponding ellipse gradient of each pixel in the annular region 300R is determined. In the present example, the gradient of gray level of pixel P is given to be (-39, 3), and its~~

~~ellipse gradient to be (0.016, -0.006), the angle θ of the two gradients of pixel P is determined by formula (4):~~

$$\cancel{\theta=\cos^{-1}\{(DX1\times Dex+DY1\times Dey)/[(DX1^2+DY1^2)\cdot(Dex^2+Dey^2)]^{1/2}\}} \quad \cancel{(8)}$$

Thus, the angle between the two gradients of pixel P is determined as 0.29 (radian). Similarly, the angles between the two gradients at other pixels P1, P2, P3, P4, P5, P6, P7, P8, etc. in the annular region 300R can be easily determined as 0.19, 0.26, 0.33, 0.18, 0.23, 0.50, 0.27, 0.42 (radian), etc., respectively.

Further, in step S45, the average of the angles between the gradient of gray level and its corresponding reference gradient for all pixels in the annular region 300R is given by formula (9):

$$A1=S1/C1 \qquad (9)$$

Where A1 denotes the average of the angles between the gradient of gray level and its corresponding reference gradient for all pixels in the annular region, S1 denotes the sum of the angles between the two gradients of all pixels in the annular region; C1 denotes the total number of pixels in the annular region.

As to the present example, the average angle between the gradient of gray level and its corresponding reference gradient of all pixels in the annular region 300R is determined as 0.59.

Returning to step S45 in Fig. 3, after the above-mentioned average of

the gradient differences has been determined, it is determined in step S45 whether said average is less than a first threshold. In the present invention, the first threshold is between 0.01 and 1.5, such as 0.61. If the calculated average is determined to be less than the first threshold, then the flow goes to step S46. If not, the flow goes to step S48. In step S48, it is determined that the rectangle 300A does not contain an image of a human face, and then, the flow ends in step S49.

In the present example, the average of the angle between the gradient of gray level and its corresponding reference gradient for all the pixels in the annular region 300R is 0.59, which is less than the first threshold. Therefore, the flow goes to step S46.

In step S46, using the weight W1 as described above, a weighted average of the angle difference and the gradient of gray level and the reference gradient for all the pixels in the annular region 300R is determined, as is given by formula (10):

$$A2=S2/C2; \qquad (10)$$

Where A2 denotes the weighted average of the angle between the two gradients for all the pixels in the annular region, C2 denotes the total number of pixels in the annular region, S2 denotes the sum of the product of the angle and the weight of gradient of gray level for all the pixels in the annular region.

In the present example, the weighted average angle between the two gradients for all pixels in the annular region is 0.66.

After the determination of the weighted average angle, it is determined in step S46 whether said weighted average angle is less than a second threshold, such as 0.68. The second threshold in the present invention is ranged from 0.01 to 1. If the weight average angle is determined to be less than the second threshold, then the flow goes to step S47, where it is determined that the rectangle 300A contains an image of a human face, then the flow ends in step S49. If the weighted average angle is not less than the second threshold, then the flow goes to step S48. In step S48, it is determined that the rectangle 300A does not contain an image of a human face, and then, the flow ends in step S49.

As to the present example, the weighted average angle for all pixels in the annular region 300R is 0.66, which is less than the second threshold, and then the flow goes to step S47. In step S47, and the region 300A is determined to contain an image of a human face. Afterwards, the flow ends in step S49.

The Second Example

Fig. 8A is a diagram showing another example of an original image to be detected. The original image 800 in Fig. 8A is produced by a digital camera. Of course, it also can be input to human face identification system by a digital device 111, such as a scanner or the like. Assume that the original image is stored in a predetermined region in the HD 107 or the RAM 109, or the like.

Fig. 8B is a diagram showing a rectangular region 800B in the original image 800, which is shown in Fig. 8A, and the location in Cartesian coordinate system, in which the X and Y axes extend in horizontal and vertical direction respectively.

As to the rectangular region 800B shown in Fig. 8B, the same flow chart shown in Fig. 4 is adopted to explain the process of determining whether the rectangular region 800B contains a human face.

Referring to Fig. 4, eye detecting device 218 determines eyes in the image, in step S41, reading means 210 reads the gray level of each pixel in the original image 800, and the sub-image determining device 219 determines rectangular regions to be identified, such as rectangular region 800B, in the original image 800. As shown in Fig. 8A, the origin of the Cartesian coordinate system is chosen to be at the upper left corner of the original image 800. Then the coordinates of each of the four corners of rectangular region 800B shown in Fig. 8B are (203, 219), (373, 219), (203, 423) and (373, 423) respectively in the Cartesian coordinate system. That is, the width and length of region 800B are 170 and 204 respectively.

Then, in step S42, the annular region 800C around the rectangular region 800B in Fig. 8B is determined by the annular region determination means 211. Referred to Fig. 8C, as to the rectangular region 800B, by scaling its width and length with a factor of 9/7 and a factor of 5/7 respectively, two rectangular regions 800C1 and 800C2 are generated.

Fig. 8C is a diagram showing the annular region 800C, which is determined for the rectangular region 800B. The annular region 800C is formed by the first rectangular region 800C1 and the second rectangular region 800C2. The coordinates of each of the four corners of first rectangular region 800C1 in this coordinate system are (170, 190), (397, 190), (179, 452) and (397, 452) respectively, while the coordinates of each of the four corners of the second rectangular region 800C2 in this coordinate system are (227,

248), (349, 248), (227, 394) and (349, 394) respectively.

Then, in step S43, the gradient of the gray level of each pixel in the annular region 800C is determined by the first determination means 212, which is (188, 6), (183, 8), (186, 10), (180, 6), (180, 6), etc. respectively. Similarly, the weight of the gradient of the gray level of each pixel in the annular region 800C is determined in the same way, which is 0.76, 0.75, 0.77, 0.77, 0.73, etc. respectively.

Further, in step S44, the reference gradient of each pixel in the annular region 800C is determined respectively by a second determination means 213, which (-0.015, -0.012), (-0.015, -0.012), (-0.015, -0.012), (-0.014, -0.012), etc. respectively. Then, the flow goes to step S45.

In step S45, the angle between the two gradients of each pixel in the annular region 800C is determined; the average angle between the two gradients for pixels in the annular region 800C can be determined in the same way as above-mentioned, which is 0.56 for the present example and is less than the first threshold. Then the flow goes to step S46.

In step S46, the weighted average angle between the two gradients for pixels in the annular region 800C is determined as 0.64, which is less than the first threshold, and the flow goes to step S47, where it is determined that the rectangular region 800B contains an image of a human face. Then, the flow ends in step S49.

Alternative Embodiment

An alternative embodiment of the present invention is illustrated below.

Fig. 9 is a diagram showing the human face determination process according to the alternative embodiment of the present invention. In Fig. 9, the same reference numeral as that in Fig. 4 denotes the same process.

Again, the original image 800 in Fig. 8A is taken as an example to illustrate determining whether a rectangular region 800B in the original image 800 contains a human face image.

First, as the same as that in Fig. 4 the process flow starts in step S40. In step S41, reading means 210 reads the gray level of each pixel in original image 800; eye detecting means 218 detect eye areas in the image; and sub-image determining means 219 determines the sub-image(s) based on the eye areas detected, such as rectangular region 800B, to be identified. Then the flow goes to step S42. In step S42, a annular region 800C is determined for rectangular region 800B, as shown in Fig. 8C. Further, the flow goes to step S43. In step S43, the gradient of the gray level at each pixel in the annular region 800C is determined, which is (188, 6), (183, 8), (186, 10), (180, 6), (180, 6), etc. respectively. The weight of the gradient of the gray level at each pixel in the annular region 800C is determined in the same way as in the previous embodiment, which is 0.76, 0.75, 0.77, 0.77, 0.73, etc. respectively. Afterwards, the flow goes to step S44. In step S44, the reference gradient of each pixel in the annular region 800C is determined, which is (-0.015, -0.012), (-0.015, -0.012), (-0.015, -0.012), (-0.015, -0.012), (-0.014, -0.012), etc. respectively.

Then, the flow goes to step S95. In step S95, the angle between the gradient of the gray level and its corresponding reference gradient at each pixel in the annular region 800C is calculated respectively e.g., in the same

way as the previous embodiment. Which is 0.64, 0.63, 0.62, 0.64, 0.64, etc. (in radian), respectively. Then, the weighted average angle between the two gradients for all pixels in the annular region 800C is calculated. While in step S95, it is determined whether the weighted average is less than a third threshold, such as 0.68. In the present invention, the third threshold is ranged from 0.01 to 1. If the weighted average angle is less than the third threshold, then the flow goes to step S96. In step S96, the rectangular region is determined to contain a human face image. If the weighted average angle is not less than the third threshold, then the flow goes to step S97, where it is determined that region 800D does not contain a human face image. Then the flow ends in step S98.

In the present embodiment, since the average of the product of the gradient difference and its corresponding gradient weight of each pixel in the annular region is 0.64, which is less than the third threshold, the flow goes to step S96 to determine that the rectangular region to be determined 800B contains a human face. Afterwards, the flow ends in step S98.

~~In the method and apparatus for human face identification, human face identification is performed by evaluating the difference between the direction of the gradient of the gray level distribution of the image to be processed and the direction of a reference gradient, such as the ellipse gradient as described above. Thus, the magnitudes of the gradient of the gray level distribution and the reference gradient do not affect the results of the evaluation.~~

~~Thus, other reference gradient may be determined for a sub-image 300A for human face identification.~~

~~For example, for each sub-image 300A determined, a~~ reference distribution~~of a human face, which can be expressed as:~~

$$\cancel{z(x,y) = -\frac{(x-x_c)^2}{a^2} - \frac{(y-y_c)^2}{b^2} + h}$$

~~is determined, wherein h is a constant, a/b equals to the ratio of the width of said sub-image to the height of said sub-image, the origin is at the center of said sub-image, (x_c, y_c) is the center of the sub-image, and z(x,y) is the gray-level value at point (x,y). Figure 10 shows an image with such a~~ reference distribution~~.~~

~~Then, an annular region 300R is generated, as explained above. After that, gradients ∇z and ∇g are calculated. ∇g is the gradient of the gray level distribution and can be calculated using any operator for discrete gradient calculation, e.g. Sobel operator. Then, the angle between vectors ∇z and ∇g is calculated.~~

~~Specifically, the angle θ between ∇z and ∇g at point (x,y) can be calculated using:~~

$$\cancel{\cos\theta = \nabla z \cdot \nabla g / (|\nabla g| \times |\nabla z|)} \qquad \text{~~and~~}$$

$$\cancel{\theta = \cos^{-1}(\nabla z \cdot \nabla g / (|\nabla g| \times |\nabla z|))}$$

~~so that the correspondence between the two vector fields ∇z and ∇g can be evaluated as explained above.~~

As a further alternative embodiment the present invention, for point (x,y), the direction of ∇g and the tangent direction of the curve

$$-\frac{(x-x_c)^2}{a^2} - \frac{(y-y_c)^2}{b^2} + h = p$$

which passes through point (x,y) and which is the contour line of z(x,y)=p, (where p is a parameter), and then evaluate the correspondence between the direction of ∇g and the tangent (or normal) direction of the curve $-(x-x_c)^2/a^2+(y-y_c)^2/b^2)+h=p$, e.g. by calculating the angle between the two directions, while p varies to generate a group of ellipses so as to cover the entire annular area 300R, and then determine whether the image contains human face(s) based on the result of the evaluation.

Note that the present invention may be applied to either a system constituted by a plurality of devices (e.g., a host computer, an interface device, a reader, a printer, and the like), or an apparatus consisting of a single equipment (e.g., a copying machine, a facsimile apparatus, or the like).

The objects of the present invention are also achieved by supplying a storage medium. The storage medium records a program code of a software program that can implement the functions of the above embodiment to the system or apparatus, and reading out and executing the program code stored in the storage medium by a computer (or a CPU or MPU) of the system or apparatus. In this case, the program code itself read out from the storage medium implements the functions of the above-mentioned embodiment, and the storage medium, which stores the program code, constitutes the present invention.

As the storage medium for supplying the program code, for example, a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM, CD-R, magnetic tape, nonvolatile memory card, ROM, and the like may be used.

The functions of the above-mentioned embodiment may be implemented not only by executing the readout program code by the computer but also by some or all of actual processing operations executed by

an OS (operating system) running on the computer on the basis of an instruction of the program code.

As can be seen from the above, the method of the present invention provides a fast approach for determining human face in a picture with a complex background, without requiring the detected picture to have a very high quality, thereby substantially eliminating the possibility of the human face in the picture being skipped over. The method allows for the precision determination of human face under different scales, orientation and lighting condition. Therefore, in accordance with the present invention, with the method, apparatus or system, the human face in a picture can be quickly and effectively determined.

The present invention includes a case where, after the program codes read from the storage medium are written in a function expansion card which is inserted into the computer or in a memory provided in a function expansion unit which is connected to the computer, CPU or the like contained in the function expansion card or unit performs a part or entire process in accordance with designations of the program codes and realizes functions of the above embodiment.

In a case where the present invention is applied to the aforesaid storage medium, the storage medium stores programs codes corresponding to the flowcharts (Fig. 4, and 9) described in the embodiments.

The embodiment explained above is specialized to determine human face, however, the present invention is not limited to determine human face, it is applicable to other determination method, for example, method to detect flaw portion on a circuit board.

As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.

Claims

1. A method for determining an object in an image having a gray-level distribution,

characterized in that said method comprises the steps of:

a)for a subset of pixels in said image, deriving a first variable from said gray-level distribution of said image,

b)for said subset of pixels, deriving a second variable from a preset reference distribution, said reference distribution being characteristic of said object;

c)evaluating the correspondence between said first variable and said second variable over the subset of pixels;

d)determining if said image contains said object based on the result of the evaluation step.

2. The method of claim 1, characterized in that said step of selecting a subset of the pixels in said image further comprises the steps of:

selecting an area in said image,

wherein pixels of said subset are in said area, and

wherein over pixels in said area the first variable as derived from the gray-level distribution of the object to be detected has better correspondence to the second variable than over pixels outside said area.

3.The method of claim 2, characterized in that it further comprises the step of:

determining a sub-image in said image; and

selecting said area in said sub-image.

4.The method of any one of claim 1-3, wherein said first variable

represents the direction of the gradient of the gray-level distribution of said image.

5.The method of any one of claim 4, wherein said second variable represents the direction of the gradient of said reference distribution.

6.The method of claim 5, wherein said gradient is a discrete gradient.

7.The method of any one of claim 5, characterized in that said evaluation step includes a statistical processing over said subset of pixels.

8.The method of claim 5, characterized in that said evaluation step includes a statistical processing over said subset of pixels and said statistical process comprises a weighting process, in which pixels with larger magnitude of gradient of the first variable have larger weights.

9.The method of claim 8, characterized in that said evaluating step includes both the statistical process comprising the weighting process and another statistical process without the weighting process and determines whether the image contains the object in accordance with the results of both of the statistical processes.

10.The method of claim 3, characterized in that said area is an annular area, which centers around the center of said sub-image.

11.The method of claim 2, wherein said subset of pixel includes all the pixels in said area.

12.The method of claim 3, characterized in that said sub-image is

determined by identifying a predetermined feature in said image.

13.The image of claim 12, characterized in that said predetermined feature includes a pair of dark areas.

14.The method of claim 13, characterized in that said sub-image is determined based on the positions of and interval between the centers of said pair of dark areas.

15.The method of claim 13, characterized in that said reference distribution is expressed by

$$z(x,y)= -(x^2/a^2+y^2/b^2)+h$$

wherein h is a constant, a/b equals to the ratio of the width of said sub-image to the height of said sub-image, and the origin is at the center of said sub-image.

16. The method of claim 5 or 15, wherein said evaluating step includes:

for all the pixels in said subset, calculating the average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference distribution;

comparing said average value with a predetermined value; and

determining that the sub-image contains said object if said average value equals to or is smaller than said predetermined value.

17. The method of claim 5 or 15, wherein said evaluating step includes:

for all the pixels in said subset, calculating a weighted average value of

the angle between the gradient of the gray-level distribution of said image and the gradient of said reference distribution, wherein a pixel having a greater magnitude of gradient of gray-level distribution has a larger weight;

comparing said weighted average value with a predetermined value; and

determining that the sub-image contains said object if said weighted average value equals to or is smaller than said predetermined value.

18. The method of claim 5 or 15, wherein said evaluating step includes:

for all the pixels in said subset, calculating the average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference distribution;

for all the pixels in said subset, calculating a weighted average value of the angle between the gradient of the gray-level distribution of said image and the gradient of said reference distribution, wherein a pixel having a greater magnitude of gradient of gray-level distribution has a larger weight;

comparing said average value with a first predetermined value;

comparing said weighted average value with a second predetermined value; and

determining that the sub-image contains said object if said average value equals to or is smaller than said first predetermined value and said weighted average value equals to or is smaller than said second predetermined value.

19. A method for determining an object in an image having a gray-level distribution, characterized in that said method comprises the steps of:

a) determining a sub-image in said image;

b) for a subset of pixels, deriving a first variable from said gray-level distribution of said image,

c) for said subset of pixels, deriving a second variable from a preset reference distribution, said reference distribution being characteristic of said

object;

d) evaluating the correspondence between said first variable and said second variable over the subset of pixels; and

e) determining if said image contains said object based on the result of the evaluation step.

20. A storage medium with a program code for object-identification in an image with a gray-level distribution stored therein, characterized in that said program code comprises:

codes for determining a sub-image in said image;

codes for deriving, for a subset of pixels, a first variable from said gray-level distribution of said image,

codes for deriving, for said subset of pixels, a second variable from a preset reference distribution, said reference distribution being characteristic of said object;

codes for evaluating the correspondence between said first variable and said second variable over the subset of pixels; and

codes for determining if said image contains said object based on the result of the evaluation step.

21. An image processing method for determining a feature portion in an image, comprising:

a read step of reading the image and a rectangle region to be determined in the image;

a setting step of setting a ~~band region~~annular region surrounding said rectangle region;

a first determination step of determining the gradient of gray level of each pixel in said ~~band region~~annular region;

a second determination step of determining a reference distribution for the

annular region and determining a ~~n ellipse for each pixel in said band region~~annular region~~, said ellipse passing through a pixel in said band region~~annular region~~, and determining the ellipse~~ reference gradient of each pixel in the ~~band region~~annular region; and

a third determination step of determining the feature portion contained in said rectangle region on the basis of the gradient of the gray level of each pixel and the ~~ellipse~~reference gradient of each pixel in said ~~band region~~annular region.

22. The image processing method according to claim 21, wherein the image contained said rectangle region to be determined is a feature portion.

23. The image processing method according to claim 21, wherein said ~~band region~~annular region is formed by a first rectangle region and a second rectangle region.

24. The image processing method according to claim 23, wherein said first rectangle region is located inside said rectangle region to be determined.

25. The image processing method according to claim 23, wherein said second rectangle region is located outside said rectangle region to be determined.

26. The image processing method according to claim 23, wherein the center of said first rectangle region lies in the same location as that of said rectangle region to be determined, and the width and the length of said first rectangle region respectively equals m times the width and the length of said rectangle region to be determined, where $0 < m < 1$.

27. The image processing method according to claim 23, wherein the center of said second rectangle region lies in the same location as that of said rectangle region to be determined, and the width and the length of said second rectangle region respectively equals n times the width and the length of said rectangle region to be determined, where $1 < n < 2$.

28. The image processing method according to claim 21, wherein said first determination step comprising the step of determining the gradient of the gray level of a pixel on the basis of the gray level of the pixels surrounding the pixel.

29. The image processing method according to claim 28, wherein the number of pixels surrounding said pixel equals k^2, where k is an integer, and $2 < k < 15$.

30. The image processing method according to claim 21, wherein ~~the center of said~~ reference distribution is determined in accordance to the annular region.~~ellipse lies in the same location as that of the inscribed ellipse in the rectangle region to be determined, and the ratio of the major axis to the minor axis of said ellipse is the same as that of the inscribed ellipse in the rectangle region to be determined.~~

31. The image processing method according to claim 21 or 30, ~~wherein~~ further comprising the step of determining the ~~difference of~~angle the gradient of the gray level of each pixel and the ~~ellipse~~ reference gradient of each pixel in said ~~band region~~annular region.

32. The image processing method according to claim 21 or 30, ~~wherein~~ further comprising the step of determining the ~~mean of the difference~~average of the angle between ~~of~~ the gradient of the gray level of each pixel and the ~~ellipse~~reference gradient of each pixel in said ~~band region~~annular region.

33. The image processing method according to claim 21 or 32, ~~wherein~~ further comprising the step of determining the weight of the gradient of the gray level of each pixel in said ~~band region~~annular region.

34. The image processing method according to claim 32, ~~wherein~~ further comprising the step of determining whether the average of the angle ~~mean of the difference of~~between the gradient of the gray level of each pixel and the ~~ellipse~~reference gradient of each pixel in said ~~band region~~annular region is less than the first threshold, where the first threshold is within the range of 0.01 to 1.5, preferably of 0.61.

35. The image processing method according to claim 34, ~~wherein~~ further comprising the step of determining whether the weighted average of the angle~~mean of the product of the difference, which is that of~~ between the gradient of the gray level of each pixel and the ~~ellipse~~reference gradient of each pixel in said ~~band~~

~~region~~annular region~~,~~ with~~and~~ the weight of the gradient of the gray level of each pixel in said ~~band region~~annular region is less than the second threshold, where the second threshold is within the range of 0.01 to 1, preferably of 0.68.

36. The image processing method according to claim 33, ~~wherein~~ further comprising the step of determining whether the weighted average of the angle ~~mean of the product of the difference, which is that of~~between the gradient of the gray level of each pixel and the reference gradient ~~of its said corresponding ellipse, and~~with the weight of the gradient of the gray level of each pixel in said ~~band region~~annular region is less than the third threshold, where the third threshold is within the range of 0.1 to 1, preferably of 0.68.

37.A human eye detection method for detecting human eye in an image, comprising:

~~the~~a reading step of reading the gray level of each pixel in the each column in the image;

~~the~~a segmenting step of segmenting each column into a plurality of intervals , and labeling each of the intervals as valley region , relay region or peak region;

~~the~~a merging step of merging the valley region of the each column and the valley region of its adjacent column, and generating an eye candidate region; and

~~the~~a determining~~ation~~ step of determining the human eye from the eye candidate regions.

38.The method according to claim 37, wherein the segmenting step includes the step of labeling each interval as one of valley region, relay region and peak region based on the gray level of each the interval in a column.

39.The method according to claim 37, wherein the segmenting step includes the step of labeling each interval as one of valley region, relay region and peak region based on the luminance value of each interval in a column.

40.The method according to claim 37, wherein the segmenting step includes

the step of labeling each interval in a column as one of valley region, relay region and peak region based on the respective gray level ratios of the interval to its two adjacent intervals.

41.The method according to claim 37, wherein the segmenting step includes the step of labeling each interval in a column as one of valley region, relay region and peak region based on the respective luminance value ratios of the interval to its two adjacent interval.

42.The method according to any one of claims 37, ~~wherein~~ further comprising the step of comparing the gray level of a segment point with that of its two adjacent intervals, and adding the segment point into the adjacent interval which has a gray level close to that of the segment point.

43.The method according to any one of claims 37, ~~wherein~~ further comprising the step of comparing the gray level of a relay region with that of its adjacent valley region and peak valley, and adding the relay region into the valley region on the basis of the gray levels of the adjacent valley region, peak region and the relay region.

44.The method according to claim 37, wherein the merging step comprising the steps of

setting each valley region in the first column of the image as a seed region respectively;

reading out the valley region of next column of the image,

determining whether the valley region can be merged into the seed region,

merging the valley region that can be merged into the seed region;

setting the non-merged valley region as a seed region; and

determining the seed region which has no further merged valley region as an eye candidate region.

45.The method according to claim 44, ~~wherein~~ further comprising the step of comparing the size of overlap of the valley region and the seed region's predicted valley region, the valley region and the seed region's predicted valley region,

46.The method according to claim 44, ~~wherein~~ further comprising the step of comparing the gray level of the valley region and that of the seed region.

47.The method according to claim 44, ~~wherein~~ further comprising the step of comparing the gray levels of the valley region, the seed region, the background gray levels of the valley region and the seed region.

48.The method according to claim 37, wherein the determining~~ation~~ step comprising the step of determining whether the gray level of the eye candidate region is less than the first threshold, where the first threshold is within the range of 100 to 200, preferably of 160.

49.The method according to claim 48, wherein the determining~~ation~~ step further comprising the step of determining whether the background gray level of the eye candidate region is bigger than the second threshold, where the second threshold is within the range of 20 to 80, preferably of 30.

50.The method according to claim 49, wherein the determining~~ation~~ step further comprising the step of determining whether the difference of background gray level of the eye candidate region and its gray level is bigger than a third threshold, where the third threshold is within the range of 5 to 120 , preferably of 20.

51. The method according to claim 50, wherein the determining~~ation~~ step further comprising the step of determining whether the ratio of the width to the height of an eye candidate region is bigger than a fourth threshold, where the fourth threshold is within the range of 1 to 5, preferably of 3.33.

52.The method according to claim 51, wherein the determining~~ation~~ step further comprising the step of determining whether the ratio of the size of an eye candidate region to that of its circum-rectangle is bigger than the fifth threshold, where the fifth threshold is within the range of 0.2 to 1 preferably of 0.4.

53. The method according to claim 14, wherein the step for determining a sub-image in said image includes:

reading the gray level of each pixel in the each column in the image;

segmenting each column into a plurality of intervals , and labeling each of the intervals as valley region , relay region or peak region;

merging the valley region of the each column and the valley region of its adjacent column, and generating a candidate dark area; and

determining a dark area.

54. The method according to claim 53, wherein the merging step comprising the steps of

setting each valley region in the first column of the image as a seed region;

reading out the valley region of the next column of the image;

determining whether the valley region can be merged into the seed region,

merging the valley region that can be merged into the seed region;

setting the valley region that cannot be merged into the seed region as a seed region; and

determining the seed region which has no further merged valley region as a dark area.

ABSTRACT OF THE INVENTION

Disclosed is a method, apparatus, and system for identifying human face in an image. The method comprises: for a subset of pixels in said image, deriving a first variable from said gray-level distribution of said image; for said subset of pixels, deriving a second variable from a preset reference distribution, said reference distribution being characteristic of said object; evaluating the correspondence between said first variable and said second variable over the subset of pixels; determining if said image contains said object based on the result of the evaluation step.

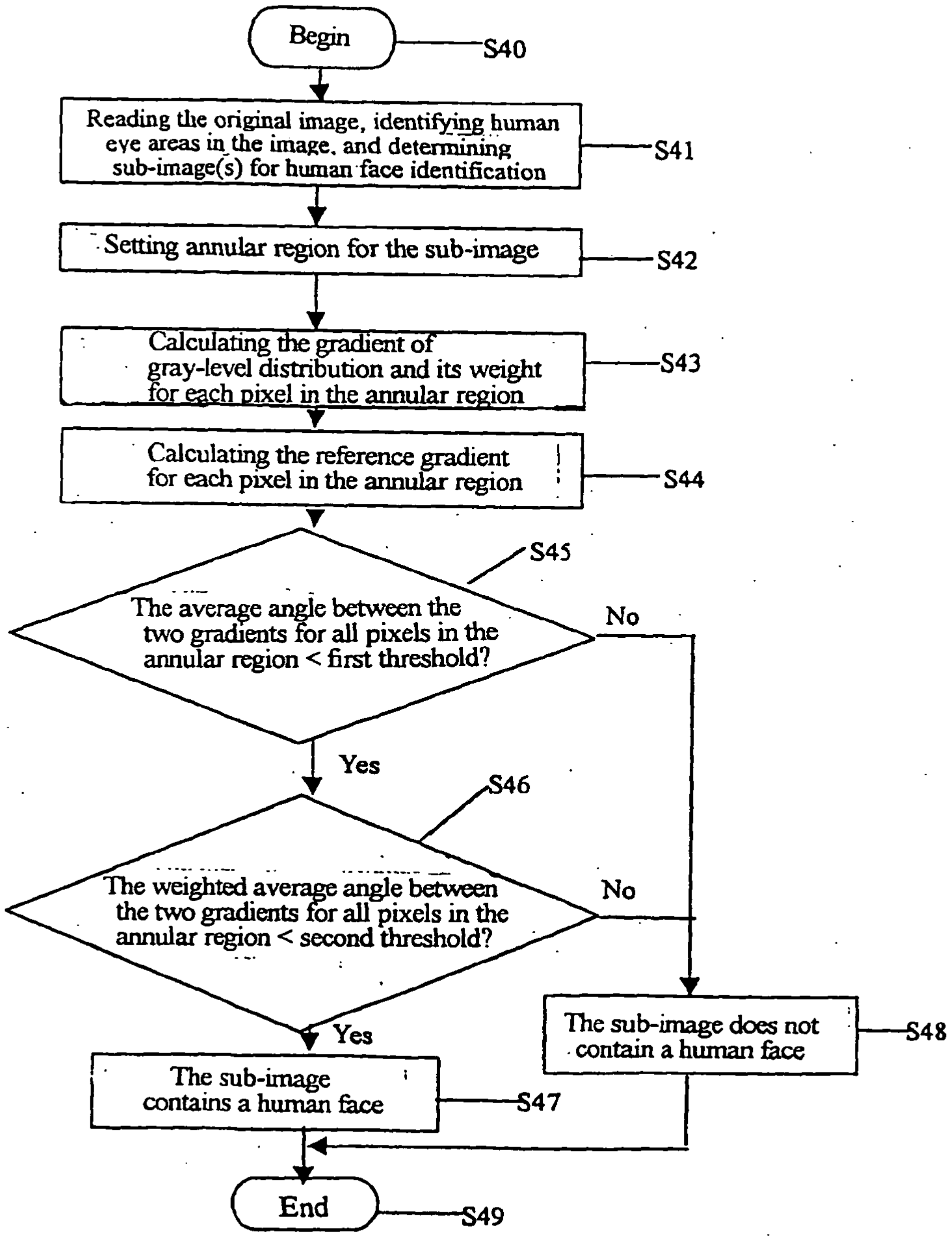

Begin
S40
Reading the original image, identifying human eye areas in the image, and determining sub-image(s) for human face identification
S41
Setting annular region for the sub-image
S42
Calculating the gradient of gray-level distribution and its weight for each pixel in the annular region
S43
Calculating the reference gradient for each pixel in the annular region
S44
S45
The average angle between the two gradients for all pixels in the annular region < first threshold?
No
Yes
S46
The weighted average angle between the two gradients for all pixels in the annular region < second threshold?
No
Yes
The sub-image does not contain a human face
S48
The sub-image contains a human face
S47
End
S49

Figure 4

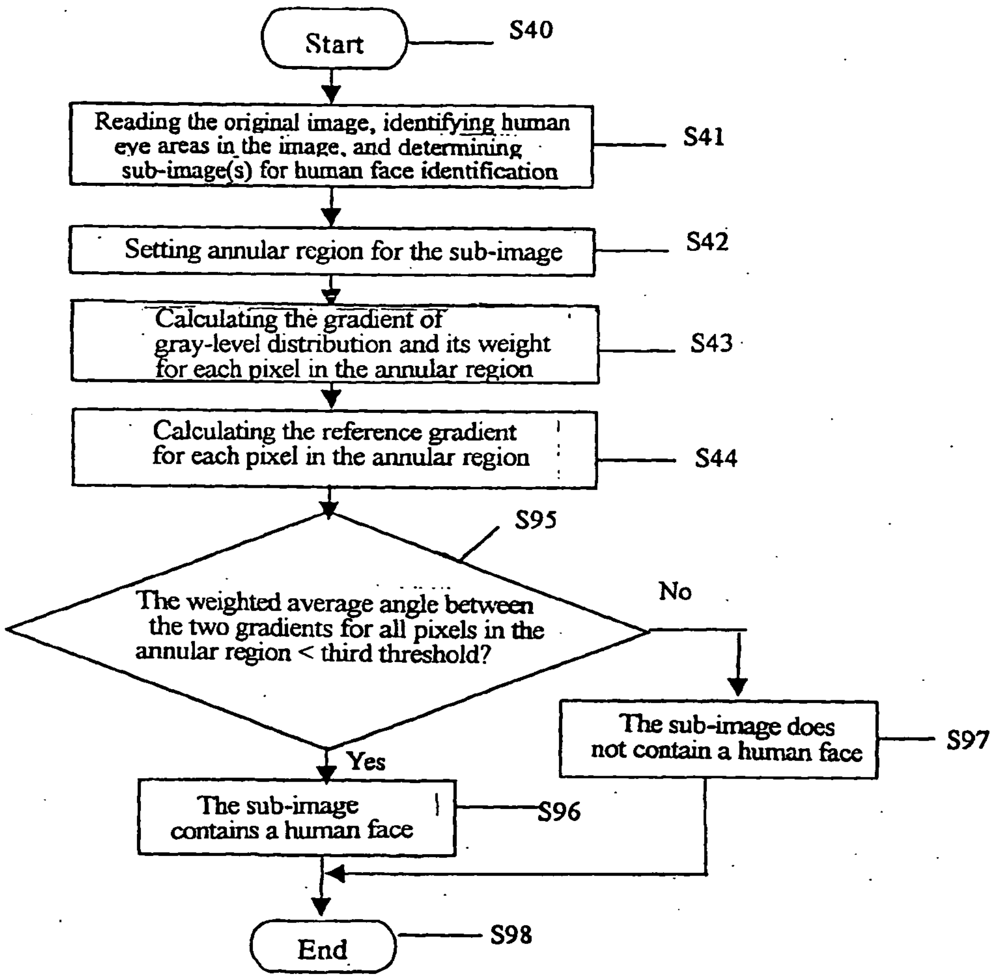

Start
S40
Reading the original image, identifying human eye areas in the image, and determining sub-image(s) for human face identification
S41
Setting annular region for the sub-image
S42
Calculating the gradient of gray-level distribution and its weight for each pixel in the annular region
S43
Calculating the reference gradient for each pixel in the annular region
S44
S95
The weighted average angle between the two gradients for all pixels in the annular region < third threshold?
No
The sub-image does not contain a human face
S97
Yes
The sub-image contains a human face
S96
End
S98

Figure 9

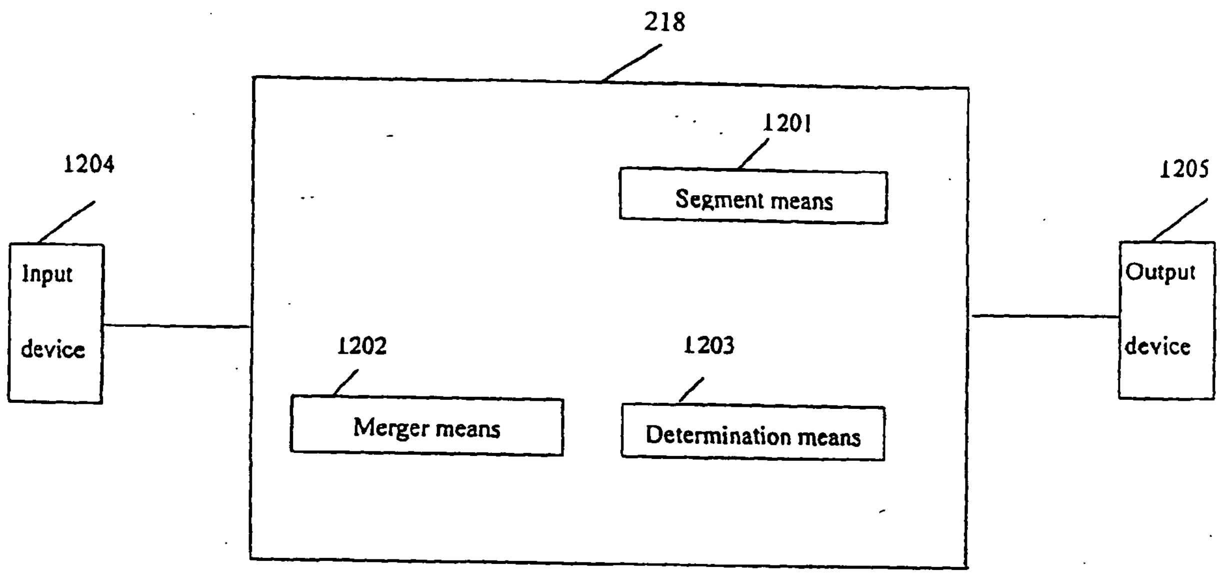

218
1201
Segment means
1204
Input
device
1205
Output
device
1202
Merger means
1203
Determination means

Fig.12

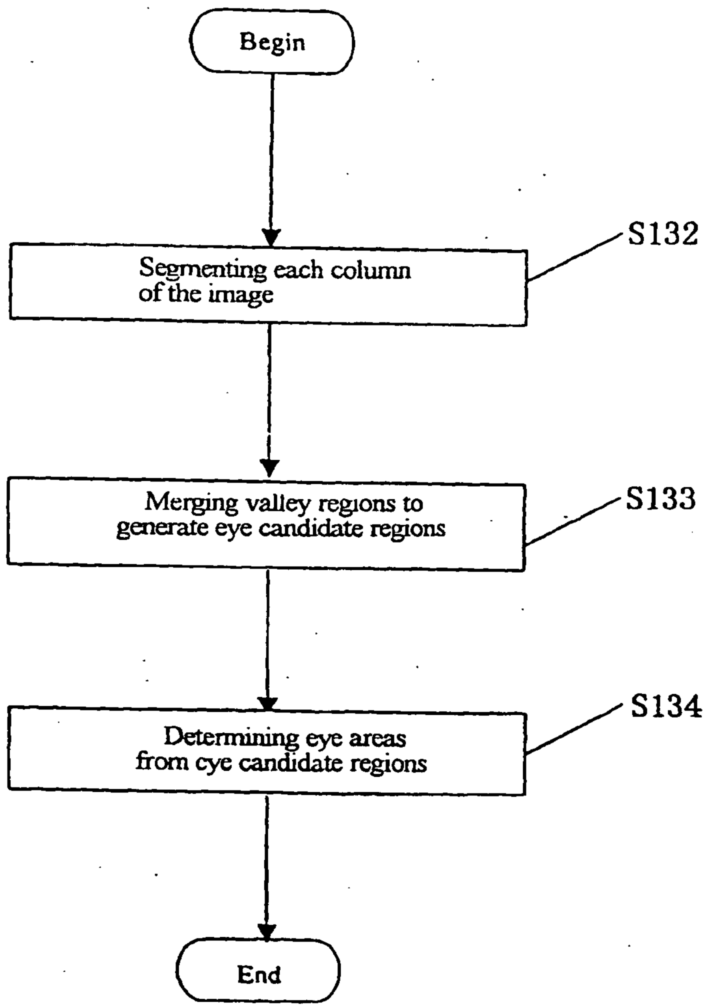

Begin
S132
Segmenting each column of the image
S133
Merging valley regions to generate eye candidate regions
S134
Determining eye areas from cye candidate regions
End

Fig.13A

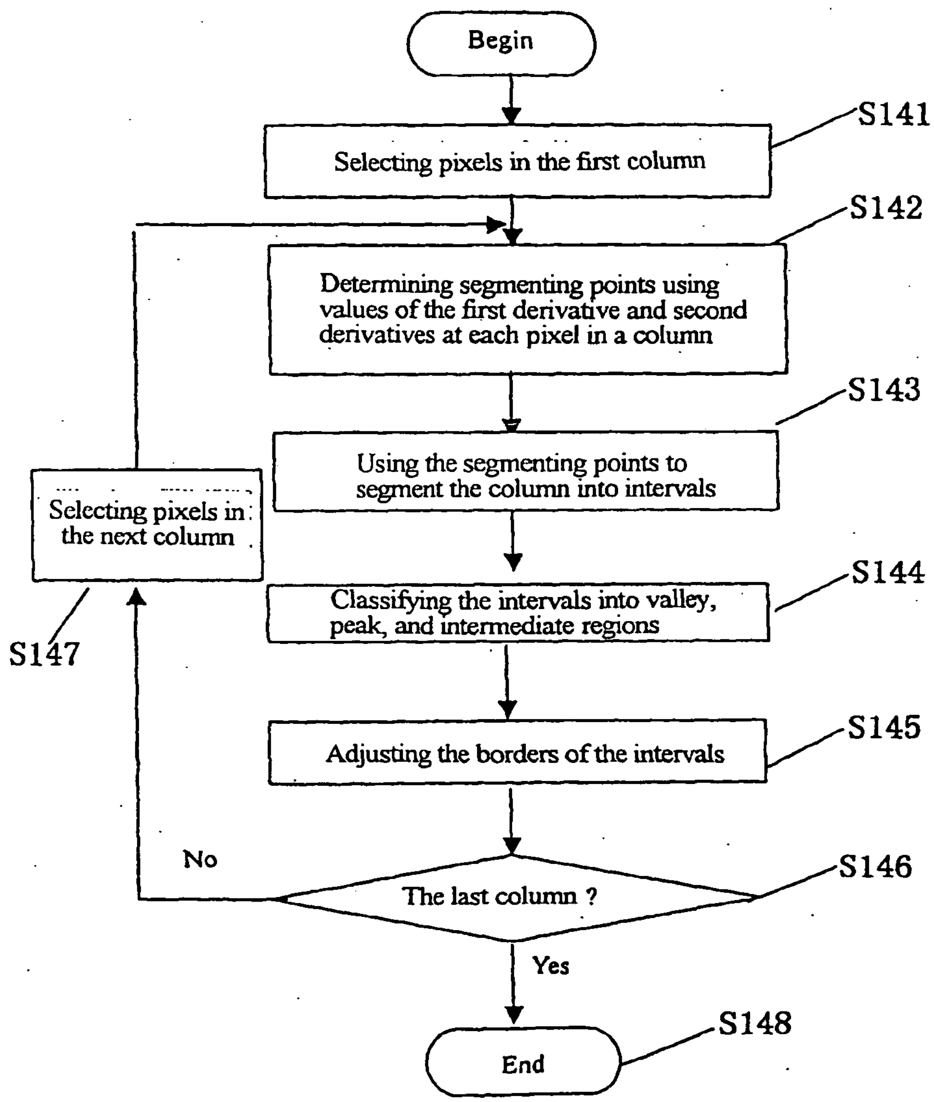

Begin
S141
Selecting pixels in the first column
S142
Determining segmenting points using values of the first derivative and second derivatives at each pixel in a column
S143
Using the segmenting points to segment the column into intervals
Selecting pixels in the next column
S147
S144
Classifying the intervals into valley, peak, and intermediate regions
S145
Adjusting the borders of the intervals
No
S146
The last column ?
Yes
S148
End

Fig.14A